REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree PhD Year 2005 Name of Author MITCHELL, D G V

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS

Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

A. Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).

B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.

C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.

D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

☑ This copy has been deposited in the Library of UCL

☐ This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

Neurocognitive Processes in Instrumental Learning, Response Reversal, Decision-making, and Psychopathy

Derek George Victor Mitchell

Department of Psychology,

University College London,

London.

Thesis submitted for

the degree of Ph.D. in Psychology,

University of London,

December, 2004.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

My sincere thanks must go to James Blair, who was, as a supervisor and mentor, incredibly enthusiastic, patient, and skilful. This work would not have begun without his encouragement and support; his ability to temper patience and guidance with the occasional good-natured admonishment greatly enhanced the quality of this thesis. I would like to thank Robert Hare and Jack Rachman, who, as my early mentors, introduced me to experimental psychology. I am grateful that their interest and enthusiasm for the field were infectious.

Special thanks go to two great friends and colleagues on both sides of the ocean, Salima Budhani and Rebecca Richell. Their friendship, discussion, and sense of humour made this experience enjoyable. NIH and UCL were excellent environments in which to complete this work. I would like to thank Cordelia Fine, Joseph Snow, Shmuel Lissek, Steve Kelly, Elizabeth Finger, Marina Nakic, Karina Peschardt, Abigail Marsh, and Linda Mah for their camaraderie and scientific discussion. I would also like to thank Patricia De La Torre, Dan Fridberg, Sarah Busis, Krystal Mondillo, and Niveen Kamel for their help with subject recruitment and coordination. In addition, I would like to acknowledge the staff and residents of HMP Wormwood Scrubs, HMP Grendon, and Broadmoor Hospital, particularly Alan Leonard and Laura Hamilton, for their help in gaining access to the appropriate resources to complete these studies.

I have been fortunate to have friends who have been so supportive and optimistic. I am particularly grateful for the tolerance and support provided by Stacey Bach, and I would also like to thank Kevin Chong for his encouragement.

Finally, I wish to thank my parents. I am indebted to my mother and father for their support and nurturing. This work was completed because of them, and for them.

Abstract

This thesis investigates the neurocognitive systems involved in emotional processing, learning, and decision-making and considers the implications of dysfunction in these systems for psychopathy. The introductory chapter reviews the anatomy and function of the orbitofrontal cortex and amygdala and discusses why impairment in these regions may contribute to psychopathy. Chapter 2 introduces an investigation of the impact of emotional material on operant behaviour that showed that individuals with psychopathy do not display appropriate emotional modulation of attention. The subsequent three chapters investigate the performance of individuals with psychopathy on forms of instrumental learning that are dissociable at both the cognitive and neural levels. Chapter 3 investigates stimulus-reinforcement learning, a form of learning thought to rely on intact amygdala functioning, and found impaired performance in individuals with psychopathy. Chapter 4 presents findings of impaired reversal learning and decision-making, which are indicative of orbitofrontal cortex dysfunction. Chapter 5 presents an investigation of conditional learning, which is a form of instrumental learning that is not dependent on the amygdala. As predicted by theories that emphasize amygdala and orbitofrontal cortex dysfunction in psychopathy, both groups performed similarly on the acquisition component of the conditional learning task, but individuals with psychopathy showed impairment on the second phase of the task when correct performance depended on reversing some of the acquired responses. Chapter 6 presents a case study that tests the analogy drawn between patients with lesions involving the orbitofrontal cortex and those with developmental psychopathy. The results show that the two disorders are characterized by sharply contrasting neurocognitive deficits. Chapter 7 reveals a functional Magnetic Resonance Imaging (fMRI) study addressing the basic question of whether discrete regions within the ventrolateral and orbitofrontal cortices are involved in signalling the need for response change and encoding the value of a reinforcer. The concluding chapter summarizes the results of the thesis and outlines future avenues of research that may arise from the work presented.

Table of Contents

List of Tables

List of Figures

Chapter 1

Psychopathy: Disordered Decisions, Emotion, and Regulation

1.1: Introduction

Psychopathy is a developmental disorder characterised by severe disruption in emotional responding and social functioning. Deficient decision-making, poor planning, and an apparent failure to be deterred by past negative consequences of their aberrant behaviour are key features of the disorder. Individuals with psychopathy typically demonstrate a range of persistently antisocial behaviours, callous affect, and manipulative interpersonal style (Hare, 1991). Patients with the disorder fail to appreciate the severity and consequences of their condition and show poor or adverse responding to traditional forms of treatment (Harris, Rice, & Cormier, 1991; Ogloff, Wong, & Greenwood, 1990; Rice, Harris, & Cormier, 1992). At the cognitive level, researchers have speculated that psychopathy is primarily the result of dysfunction within systems mediating fear (Gray, 1987; Lykken, 1995), response modulation (Newman, 1998), decision-making (Bechara, Damasio, Damasio, & Anderson, 1994; Damasio, 1994), empathy (Blair, 1995; Blair, Jones, Clark, & Smith, 1997), and general affective processing (Hare, 1998). However, on an individual basis, each explanation addresses only a limited number of clinical phenomena, and despite some overlap, the theories do not explain the spectrum of empirical findings associated with the disorder. For example, hypotheses that focus on empathy dysfunction explain problems identifying facial expressions in this population (Blair, Colledge, Murray, & Mitchell, 2001), but do not address difficulties in extinguishing a previously rewarded response (Newman, Patterson, & Kosson, 1987). Attempts to develop an overarching explanation of the seemingly disparate series of findings have led researchers to examine neural substrates. Two structures that have

received particular attention are the orbital and ventral regions of the frontal cortex (Blumer & Benson, 1975; Damasio, 1994; LaPierre, Braun, & Hodgins, 1995) and the amygdala (Blair, Morris, Frith, Perrett, & Dolan, 1999; Patrick, 1994).

This thesis is an investigation of the potential role that neurocognitive dysfunction might play in the presentation of psychopathy. In this chapter, two key neural regions implicated in psychopathy, the ventrolateral and orbital frontal cortices and amygdala, will be discussed in separate sections. First, empirical and anecdotal accounts of changes in personality following damage to these brain regions are reviewed to place recent experimental investigations in context. Second, a discussion of the dominant theories of function and a brief review of data from animal, patient, and human neuroimaging studies relevant to these theories will be presented. The theories and scientific results concerning function are discussed with reference to findings involving psychopathic individuals. The chapter concludes with an overview of the key questions that instigated the experimental work presented in this thesis.

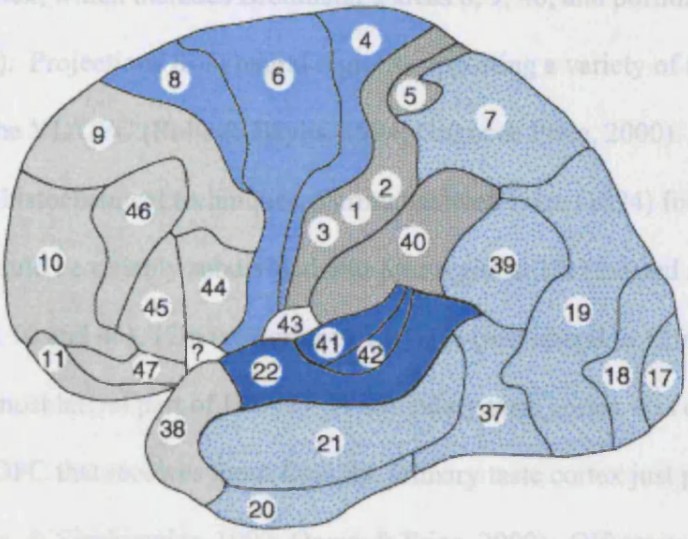
1.2: The Orbital, Medial, and Ventrolateral Regions of the Frontal Cortex

1.2.1: Anatomy

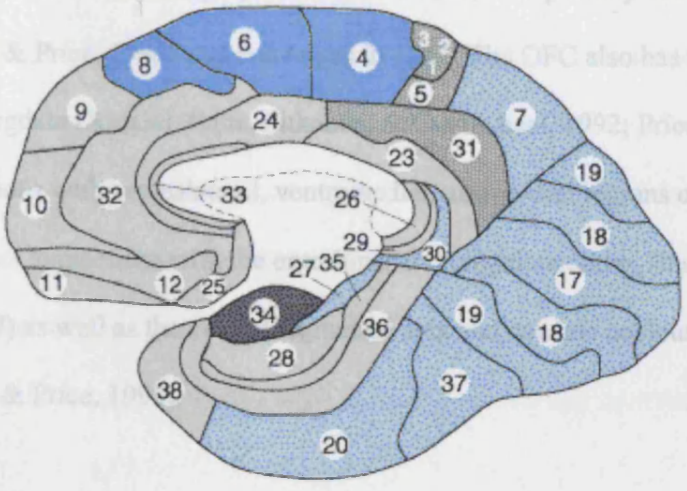
The primate orbitofrontal cortex (OFC) is defined as the areas occupying the ventral surface of the frontal part of the brain, extending laterally to include Brodmann's areas 10, 11, and 13.

Figure 1.1: Brodmann's areas of the human brain

A) Sagittal view of lateral portions of the left hemisphere



B) Sagittal view of the medial wall of the right hemisphere



The "ventrolateral prefrontal cortex" (VLPFC) corresponds to Brodmann's area 47 in humans, and area 12 in macaque monkeys (Petrides & Pandya, 2002); the term "OFC" will be used here to refer to Brodmann's areas 10, 11 anteriorally, 12 caudally, 13 caudally, and 14 medially (Carmichael & Price, 1995; Petrides & Pandya, 2002; Rolls,

2004). Regions 11 and 13, not seen in Figure 1.1, are on the ventral surface between areas 47 and 14 medially. Thus, the VLPFC and OFC (VL/OFC) refer collectively to medial, ventral, and lateral areas of the frontal cortex to the exclusion of dorsolateral prefrontal cortex, which includes Brodmann's areas 8, 9, 46, and portions of area 10 (Fuster, 1997). Projections from neural regions supporting a variety of sensory modalities converge in the VL/OFC (Rolls & Baylis, 1994; Ongur & Price, 2000). Using sophisticated histochemical techniques, Carmichael and Price (1994) found the VLPFC (BA 47/12) could be reliably subdivided into four regions: 12r (located anteriorally bordering BA 10 and 46), 12m (posterior to 12r), 12l (just lateral to 12m bordering 45) and 12s (the most lateral part of BA 47). A secondary taste cortex was discovered in the caudolateral OFC that receives input from the primary taste cortex just posterior to BA47 (Rolls, Yaxley, & Sienkiewicz, 1990; Ongur & Price, 2000). Olfactory stimulation is conveyed to the OFC via the olfactory bulb and olfactory cortex (Rolls & Baylis, 1994). Finally, tactile information is projected to the OFC via the primary somatosensory cortex (Carmichael & Price, 1995) and converges in 12m. The OFC also has strong connections with the amygdala (Amaral, Price, Pitkanen, & Carmichael, 1992; Price, 2003), which has rich connections with ventrolateral, ventromedial, and medial regions of PFC (Ongur & Price, 2000). Connections with the entorhinal and cingulate cortex (Insausti, Amaral, & Cowan, 1987) as well as the ventral tegmental area and caudate nucleus are also prevalent (Carmichael & Price, 1995).

1.2.2: Clinical correlates of OFC dysfunction

The suggestion that OFC dysfunction contributes to psychopathy is based on clinical observations and empirical reports of patients who have acquired lesions to this

area. One of the earliest reports of behavioural problems following frontal lobe damage involved a 19th century rail worker who, following an accident, showed a dramatic change in personality and emotional reactivity (Harlow, 1868). Such patients often undergo a marked alteration in their personality that includes behavioural disinhibition, an inability to form realistic plans, and a failure to adjust their course in response to negative prospects or consequences (Blumer & Benson, 1975; Damasio, 1994; Rolls, 1996). Clinical accounts of patients with acquired lesions describe previously high-functioning individuals who sustain a head injury and subsequently suffer from a litany of social and behavioural problems including difficulties maintaining employment, financial recklessness, inappropriate sexual behaviour, aggression, and unrealistic planning (Damasio, 1994; Hornak et al., 2003; Hornak, Rolls, & Wade, 1996; Rolls, Hornak, Wade, & McGrath, 1994). The apparent similarities between patients with acquired OFC damage and developmental psychopathic individuals have prompted some to refer to the former as "pseudo psychopaths" (Blumer & Benson, 1975) or "acquired sociopaths" (Damasio, 1994). The changes in personality and affect often manifest in poor decision-making and aggression. Research involving neuroimaging and animal models supplement the clinical accounts and implicate the OFC in skills related to adaptive social and emotional functioning such as expression recognition, decision-making, response change or extinction, and reward processing. In the sections that follow, experiments that provide evidence for dysfunction in each of these domains will be reviewed.

1.2.3: Neurocognitive investigations and theories related to OFC function

In line with suggestions of reduced empathy (Barrash, Tranel, & Anderson, 2000), patients with OFC lesions show multimodal emotional expression recognition deficits. The OFC cortex has anatomical connections with regions that are involved in

representing facial and vocal expressions, such as the temporal lobe (Rolls, 1999). Single-unit recording studies reveal neuronal activity in the ventromedial prefrontal cortex in the presence of negative, but not positive or neutral, visual stimuli (Kawasaki et al., 2001). Neurons also exist within the ventral prefrontal cortex that are selectively responsive to primate vocalizations (Romanski & Goldman-Rakic, 2002). Imaging data have shown greater activation of orbital regions for emotional rather than neutral vocal expressions (Morris, Scott, & Dolan, 1999; Phillips et al., 1998). Neuroimaging data suggest that depictions of emotional facial expressions activate regions of the VL/OFC (Blair et al., 1999; Kesler/West et al., 2001; Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998). For example, the VL/OFC shows greater activation for positive versus neutral expressions (Dolan et al., 1996) as well as greater activation to angry than to sad facial expressions (Blair et al., 1999). Studies with patients provide at least partial support for the imaging results. Thus, individuals with lesions to the area show difficulties identifying vocal and facial expressions of emotion (Hornak et al., 1996). However, a later study failed to replicate these findings in patients with bilateral OFC damage (Hornak et al., 2003); instead, vocal affect recognition impairments were evident in patients with lesions outside of the OFC, such as the anterior cingulate and more dorsal medial regions of the prefrontal cortex (BA 9; Hornak et al., 2003).

The theoretical formulation of OFC function that best accounts for data concerning deficient recognition and responding to vocal and facial expressions of emotion is the Social Response Reversal theory (Blair & Cipolotti, 2000). The Social Response Reversal explanation emphasizes the communicative role of facial expressions and suggests that the OFC responds to negative emotional expressions to monitor, assess, and adjust behaviour for ongoing and future interactions (Blair & Cipolotti, 2000). The theory extends this activation to include not only negative cues, but also situations that

have become associated with the expression of these cues. Accordingly, significant lateral OFC activation is reported when participants were asked to process scenarios that were likely to result in anger in others (Berthoz, Armony, Blair, & Dolan, 2002). Individuals who have sustained damage to the OFC are comparably insensitive to these social cues, and so cannot incorporate signals from this system to adapt their interactions to changes in their social environment.

The VL/OFC is also involved in making decisions when the future consequences of responding are unknown. Two paradigms have been recently devised to investigate decision-making: the Iowa Gambling Task (Bechara et al., 1994; Bechara, Damasio, Damasio, & Lee, 1999), and the Cambridge Gambling Task (Rogers, et al., 1999a; 1999b). Both are gambling tasks in which the reward outcomes cannot be precisely determined, but rather, are probabilistically estimated. The lack of certainty regarding the outcome is designed to mimic real-life decision-making. In the Iowa Gambling Task, participants select from four decks of cards each with a different rate and value of reinforcement. Unknown to the participant, two of the decks are disadvantageous; selecting from the decks will render high reward, but even higher punishment resulting in a net loss. Two of the decks are advantageous; consistently selecting from these decks will result in a net gain. A Positron Emission Tomography (PET) study has shown ventrolateral and OFC activation (BA 11 and 47) during the performance of the Iowa Gambling Task, but also activity in dorsolateral prefrontal cortex and anterior cingulate (Ernst et al., 2002). Patients with OFC damage persist in making risky, disadvantageous responses while healthy individuals learn to avoid such selections (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Bechara et al., 1994; Bechara et al., 1999). Furthermore, skin conductance responses generated in anticipation of disadvantageous choices were significantly lower in patients with acquired lesions.

In the Cambridge Gambling Task, participants make probabilistic judgements about whether a token is hidden under a red or blue box (Rogers, et al., 1999a). On each trial, the ratio of red to blue boxes varies, and can be 9:1, 8:2, 7:3, or 6:4 each time. There are two variants of the task. In one variant, the reinforcement levels are pre-specified (choosing the colour with the lowest probability always yields a higher reward, but also the potential for an equally high punishment). In a second variant, potential wagers are displayed on the screen serially in ascending or descending order of magnitude. The participant is free to select the wager when it appears on the screen. The task therefore synthesizes the conflict that accompanies risky decisions when dealing with uncertain outcomes. Selecting a high wager early for a low probability reward or a low wager early for a high probability reward is considered an indication of impulsivity. A PET study of the Cambridge Gambling Task reports activity in the ventral regions of OFC including Brodmann's areas 10, 11, and 47 (Rogers, et al., 1999b). Patients with lesions to the OFC show decision-making performance decrements relative to comparison individuals (Mavaddat, Kirkpatrick, Rogers, & Sahakian, 2000; Rogers et al., 1999a); however, a recent study failed to replicate this result in a sample of patients with unilateral OFC lesions (Manes et al., 2002).

Damasio (1994) suggests that the impaired decision-making of the kind seen in gambling tasks results when "somatic markers" fail to provide the affective colouring that normally guides behaviour in healthy individuals. According to the somatic marker hypothesis, the OFC is involved in integrating bodily feedback to influence decision-making (Damasio, 1994). The theory suggests that feedback from bodily states are conveyed to somatosensory structures which help precipitate judgements about the value of a particular response option. This link is thought to be possible in two ways. When this occurs via actual bodily feedback ("body loop"), somatic markers are conveyed from

the periphery to the somatosensory structures where they influence behaviour. Alternatively, the same information can be conveyed via virtual bodily feedback (an "as-if body loop") whereby the body is bypassed and instead representations of bodily responding formed on the basis of previously conditioned associations would influence behaviour via the somatosensory structures. Effectively, the somatic marker labels a particular response option as either advantageous or disadvantageous, thereby providing an automated index of the incentive value of the response. The authors suggest that the generation of an autonomic response (or somatic marker) when considering response options is crucial for adaptive decision-making and behavioural control; the failure to do so contributes to the pattern of maladaptive and often antisocial behaviour observed following OFC damage.

Perhaps related to the results concerning decision-making, it has been suggested that the OFC is responsible for adapting behaviour to changes in reinforcement contingencies (Rolls, 1996). Electrophysiological studies with non-human primates provide evidence for how reinforcement values may be flexibly coded in the OFC to facilitate relearning. For example, the involvement of the OFC in the motivational control of goal-directed behaviour is reinforced by findings that neurons within this region discriminate between reinforcers based on the animal's relative preference among available options (Tremblay & Schultz, 1999). Neurons within the OFC have been identified that respond selectively to a particular food and cease responding when the animal is fed to satiety (Critchley & Rolls, 1996). However, the OFC is not crucial for all forms of motivational responding. For example, monkeys with lesions to the OFC fail to alter behaviour to a reversal of reinforcement valence during visual discrimination, but can learn to shift attention to a previously irrelevant dimension of a compound stimulus and respond accordingly for reward (a process called "attentional set-shifting;" Dias et al.,

1996). Conversely, lesions to the dorsolateral prefrontal cortex result in impaired attentional set-shifting, but intact response-reversal (Dias, et al. 1996). Imaging studies also implicate the OFC in response change. For example, regions of VL/OFC (BA 47) are activated before response change in a probabilistic response-reversal paradigm in which selecting the correct stimulus results in reward on most but not all trials (Cools, Clark, Owen, & Robbins, 2002; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). In classic response-reversal paradigms involving humans, participants learn that responding to one of two simultaneously presented objects results in reward, and that responding to the other results in punishment. When the reward contingencies of the two stimuli are reversed unexpectedly, patients with OFC lesions show impaired ability to cease making a previously rewarding response in favour of a previously punishing response relative to patients with lesions to other regions (Berlin, Rolls, & Kischka, 2004; Fellows & Farah, 2003; Hornak et al., 2004; Rolls et al., 1994).

Based on findings related to response change, Rolls has suggested that the primary role of the OFC is to allow for the modification of behaviour in order to carry out effective reward and punishment-related responses, particularly when reinforcement contingencies change (1996; 2000). Thus, the "response-reversal" explanation suggests that this system is crucial in altering previously acquired stimulus-reward associations when they are no longer appropriate. Interestingly, Rolls' (1996; 2000) conceptualisation can also explain the impaired performance of patients with OFC lesions on the Iowa Gambling Task. In this task, all initial responses are rewarding. Indeed, the disadvantageous decks initially appear to be preferable because they offer the highest magnitude of reward. It becomes apparent that the decks with the highest magnitude of reward also offer an even higher magnitude of punishment only after a participant has made several selections from a particular deck. By then, an initial stimulus-reward

association had has been made, and an adjustment in responding is required to show adequate performance on the task. Accordingly, a response-reversal deficit may result in difficulties shifting responding from the ostensibly high magnitude reward decks to the truly advantageous decks. Rolls (1999) suggests that difficulties altering behaviour to non-reward can explain both the behavioural problems (tantrums and antisocial behaviour) and personality characteristics (emotional lability) seen in patients with OFC lesions.

1.2.4: Dysfunction within the VL/OFC and aggressive disorders

Taken together, the theories described provide reasonable accounts of the data concerning expression recognition, decision-making, and response change. However, it is not immediately apparent how these deficits would lead to heightened levels of aggression. Investigations have indeed linked metabolic and structural abnormalities in the prefrontal regions to aggression in general, and more specifically, to reactive aggression (best characterised by a frustrated temper tantrum rather than goal-directed instrumental aggression). Thus, Raine and his colleagues have used PET techniques to reveal reduced frontal glucose metabolism in violent offenders (Raine, Buchsbaum, & LaCasse, 1997; Raine et al., 1994; Raine et al., 1998). A volumetric analysis of the frontal lobes using MRI revealed an 11% reduction in grey matter volume in individuals with antisocial personality disorder (Raine, Lencz, Bihrle, LaCasse, & Colletti, 2000). Soderstrom and his colleagues (2000) report reduced regional cerebral blood flow in the frontal lobes in a group of non-psychotic violent offenders in the absence of structural damage. More pertinently, studies that have explicitly investigated the incidence of reactive aggression in their sample report a particular link between reduced prefrontal activity and reactive aggression (Volkow & Tancredi, 1987; Volkow et al., 1995). Raine

and his colleagues (1998) show a dissociation between reactive (referred to as "affective") and instrumental (referred to as "predatory") offenders. They found that while reactive offenders showed reduced left and right prefrontal functioning, instrumental offenders showed prefrontal functioning similar to the comparison group.

Recently, a unified model, the Integrated Emotion System (IES) was developed that describes how the neurocognitive deficits found in patients with OFC lesions can lead to reactive aggression (Blair, 2004). This conceptualisation draws on data from studies with non-human mammals suggesting that a neural circuit involving the OFC exists that plays a role in mediating reactive aggression (Panksepp, 1998). Blair (2004) suggests that in humans, the OFC modulates a subcortical system involved in reactive aggression as a function of the responding of the response-reversal (RR; Rolls, 2000) and social response-reversal (SRR; Blair & Cipolotti, 2000) systems. Reactive aggression is thought to be regulated by neural circuitry that humans share with other mammalian species and that is crucial for an organism's response to threat in the environment (Gregg & Siegel, 2001; Panksepp, 1971). This circuit runs from the medial amygdaloidal areas and through the stria terminalis to the medial hypothalamus, and down to the periaqueductal grey (PAG). The OFC regulates the responsiveness of this threat system through projections to the medial hypothalamus and medial nucleus of the amygdala (Gregg & Siegel, 2001; Panksepp, 1998). Effectively, function in the OFC can either diminish or augment subcortical "threat system" responding, thereby increasing or decreasing the probability of reactive aggression.

The RR system is potentially implicated in modulating the subcortical threat systems related to reactive aggression due to its role in providing flexible responding to changes in the reinforcement value of stimuli (Rolls, 2000). Frustration, thought to be one of the precursors to reactive aggression (Berkowitz, 1993), results when a behaviour

initiated to obtain reinforcement does not result in the expected reward. Consequently, deficits in the system involved in encoding expectations, violations of these expectations, or the alteration of behaviour in response to these violations may lead to increased frustration and subsequently, a higher risk for reactive aggression.

The SRR system is implicated in the modulation of the subcortical threat system because of its role in modifying behaviour as a consequence of negative social cues in the environment. Blair (2004) suggests that the activation of the SRR also occurs as a function of the agent's position in the social hierarchy. For example, in "sham" rage experiments, animals who have undergone neural stimulation will display reactive aggression to more submissive animals, but not to those in more dominant positions within the hierarchy (Alexander & Perachio, 1973). The implication is that in humans, displays of social disapproval or anger from super-ordinates will result in avoidance, the elicitation of distress cues, or other submissive displays. In contrast, the same aggressive cues exhibited by subordinates will result in confrontational behaviour. Individuals with dysfunction in the OFC affecting the SRR system will show poor modulation of subcortical threat system due to insensitivity to displays of social disapproval and poor processing of the status of individuals engaged in the social interaction. At present, it is unclear to what extent the neuroanatomical systems involved in RR and SRR overlap. Although probably anatomically similar, the primary difference between the RR and the SRR system is that the latter incorporates information with respect to status within the dominance hierarchy (a social construct) to modulate responding. In contrast, the RR system alters behaviour purely on the basis of the incentive value of the response without respect to social context. Putting one's feet up on a desk can be a rewarding experiencing and would therefore not activate the RR system. However, recognizing the impropriety of doing so on the department head's desk would be facilitated by the SRR system.

It has been suggested that OFC dysfunction contributes to developmental psychopathy (Anderson et al., 1999; LaPierre et al., 1995; Mitchell, Colledge, Leonard, & Blair, 2002). In addition to the clinical similarities, psychopathic individuals have shown difficulties in laboratory tasks that are thought to be sensitive to OFC dysfunction. For example, LaPierre and her colleagues (1995) report response-reversal deficits on a go/no-go task in which participants first acquired a prepotent response, and then were asked to withhold that response on a second phase of the task (LaPierre et al., 1995). Participants also showed difficulties on two other proposed indices of frontal functioning: olfactory discrimination, and qualitative errors on the Porteus Maze Task. Individuals with psychopathy have also shown difficulties with extinguishing a previously rewarded response on the One-Pack Card Playing Task (Newman et al., 1987). At odds with these findings, one study has shown that individuals with psychopathy did not exhibit decision-making deficits on the Iowa Gambling Task (Schmitt, Brinkley, & Newman, 1999); however, as the authors note, this study contained important procedural differences to the ones that Bechara and his colleagues implemented (Bechara et al., 1999).

Given the apparent similarities in the presentation of developmental psychopathy and acquired sociopathy, it is not surprising that that the two are surmised to have a common aetiology. However, a crucial difference exists in the type of aggression exhibited by these two patient groups. The type of aggression most often seen in individuals with lesions involving the OFC is reactive in nature (Anderson et al., 1999; Anderson, Damasio, Tranel, & Damasio, 1997; Grafman, Schwab, Warden, Pridgen, & Brown, 1996). Reactive aggression is described as a non-goal-directed response to frustration or a perceived threat. In contrast, instrumental aggression is goal-directed and usually aimed at material gain (Berkowitz, 1993). Although also showing high rates of reactive aggression, instrumental aggression has been identified as a common, and more

defining feature of psychopathy (Cornell et al., 1996; Williamson, Hare, & Wong, 1987). In contrast, reactive but not instrumental aggression is evident in individuals with acquired sociopathy (Blair, 2004). Unlike reactive aggression, instrumental aggression is not necessarily accompanied by high emotional intensity. It has been suggested that while OFC dysfunction may contribute to some of the features of the disorder, a key contribution to psychopathy is abnormality within the amygdala (Blair, 2003a, 2003b).

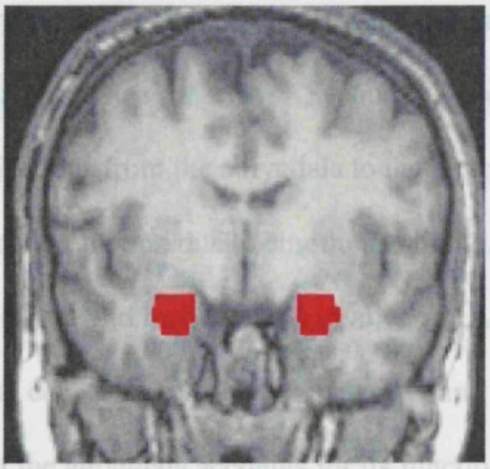
1.3: The Amygdala

1.3.1: Anatomy

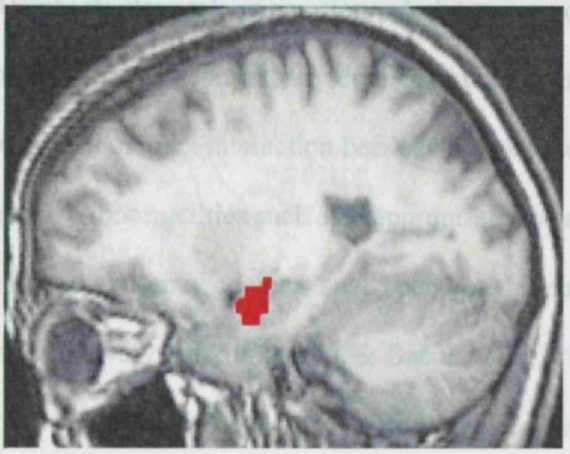
The established research suggests that the amygdala plays a pivotal role in emotion. Located bilaterally deep within the temporal lobes, the cluster of nuclei now referred collectively as the "amygdala" was first so-named by Burdach in 1819 (Aggleton, 2000).

Figure 1.2: The amygdala

A) Coronal slice taken from an MRI structural scan showing the amygdala in red located bilaterally in the temporal lobes



B) Sagittal view showing the left amygdala.



The amygdala has connections to subcortical and cortical structures involved in decision-making, reward processing, attention, memory, and sensory representation. Some of the major subcortical connections of the amygdala include the nucleus accumbens, olfactory structures, regions of the hypothalamus, thalamus and periaqueductal grey as well as the ventral tegmentum (Aggleton, 2000; Amaral et al., 1992). The cortical connections of the amygdala include regions of the frontal cortex (orbitofrontal, dorsolateral, medial), the insula, and anterior cingulate. The connectivity of the amygdala has been described as involving three main systems (Price, 2003). First, a system of inputs predominantly associated with the forebrain provides sensory information originating in the olfactory, taste, visceral pathways, posterior thalamus, and sensory association cortex pathways. Second, a system characterised as a series of reciprocal projections from the amygdala to the hypothalamus and brain stem, which modulates visceral function due to emotional stimuli. Third, a system involved in emotional behaviour and mood, which involves the ventromedial frontal, rostral insular, rostral temporal cortical areas, medial thalamus, and ventromedial basal ganglia. The rich connectivity of the amygdala accounts for its functional diversity, and its ability to affect a range of behaviours and cognitive processes.

1.3.2: Clinical correlates of amygdala dysfunction

Despite its clear role in emotion, a link between damage to the amygdala and aggression has not been consistently demonstrated. This failure to establish a link may be indicative of the complex interaction between developmental factors and the functions of the amygdala. Abnormalities such as inappropriate social interactions, increased withdrawal, and abnormal contacts have been noted in monkeys with early lesions to the amygdala (Malkova, Mishkin, Suomi, & Bachelvalier, 1997). Kluver-Bucy syndrome, characterised by reduced affect, hyper-orality, hyper-sexuality and reduced fear of threatening objects, was first noted by the researchers after which it is named in monkeys with lesions involving the amygdala (Kluver & Bucy, 1997). However, the earlier lesion studies inducing Kluver-Bucy syndrome have been criticised. The lesion techniques used often destroy not only the amygdala, but also fibres passing through the amygdala and adjacent neural regions. Studies utilizing more advanced and selective lesion techniques using ibotenic acid, which destroys neurons of the amygdala but spares the fibres passing through it, have suggested a more complex role of the amygdala in social behaviour (Prather et al., 2001). Specifically, adult and neonatal monkeys with bilateral amygdala lesions show reduced fear of threatening stimuli such as snakes. However, adult monkeys with amygdala lesions also show heightened affiliative social behaviour (Emery et al., 2001). Monkeys receiving lesions at two weeks of age can generate appropriate facial expressions and play behaviour, but show heightened fear behaviour during dyadic social situations (Prather et al., 2001). It is suggested therefore, that the amygdala plays a role in evaluating the "safety" of a social situation, but is not crucial for social cognition more generally.

Abnormalities in social behaviour have been noted in humans following lesions to the amygdala, and these abnormalities tend to be more severe with earlier onset (Fine &

Blair, 2000; Malkova et al., 1997b). Kluver-Bucy syndrome has been associated with amygdala lesions in humans (Hayman, Rexer, Pavol, Strite, & Meyers, 1998; Yoneoka et al., 2004), as has reductions in the experience of anger and fear (Sprengelmeyer et al., 1999). The clinical presentation of patients with amygdala lesions differs from that produced by OFC lesions in type and consistency. For example, anomalies in amygdala function have been associated with depression (Drevets et al., 1992), Alzheimer's disease (Chow & Cummings, 2000), schizophrenia (Fudge, Powers, Haber, & Caine, 1998), and autism spectrum disorders (Baron-Cohen et al., 2000). Bilateral amygdalectomies have been reported to decrease aggressive behaviour in 70 to 76% of cases (Ramamurthi, 1988). However, severe amygdalar atrophy has also been associated with aggression in a subgroup of patients with temporal lobe epilepsy (van Elst, Woermann, Lemieux, Thompson, & Trimble, 2000). It would therefore appear that damage to the amygdala might either increase or decrease the probability of aggression depending perhaps upon contextual or other parameters. At present, the specific parameters are unclear.

1.3.3: Neurocognitive investigations and theories related to amygdala function

Empirical data implicate the amygdala in emotional expression processing, aversive conditioning, instrumental learning, and the modulation of attention and memory. Patients with lesions to the amygdala have been shown to be impaired in the recognition of emotional facial expressions, particularly fear (Adolphs, Tranel, & Damasio, 2001; Adolphs, Tranel, Damasio, & Damasio, 1994; Adolphs et al., 1999; Anderson & Phelps, 1998). Sophie Scott and her colleagues have reported impaired recognition of vocal expressions of fear and anger in a patient with bilateral amygdala lesions (Scott et al., 1997). However, other studies involving patients with unilateral and bilateral lesions have not found significant differences relative to comparison groups

(Adolphs & Tranel, 1999; Adolphs et al., 2001; Anderson & Phelps, 1998; Anderson & Phelps, 1998). Although there have been suggestions that amygdala dysfunction results in a generalized recognition impairment (Rapcsak et al., 2000), these studies often involve patients with lesions extending beyond the amygdala. Imaging studies suggest a more specialized role for the region; modulated activity has been noted particularly with fearful (Breiter et al., 1996; Morris et al., 1996; Phillips et al., 1998; Phillips et al., 1997), sad (Blair et al., 1999) and happy (Breiter et al., 1996) facial expressions.

One of the earliest and most consistent findings in the literature regarding the amygdala is its involvement in aversive and appetitive conditioning. Animal studies show that lesions of the central nucleus of the amygdala impair processing of conditioned stimuli (CS), unconditioned stimuli (US), and unconditioned stimulus-response (UR) associations for both appetitive and aversive stimuli. For appetitive stimuli, lesions of the amygdala will result in impaired conditioning as indexed by several behavioural analogues: approach or orienting behaviours, reinforcer devaluation, pavlovian instrumental transfer, and impaired second order conditioning or second order instrumental responding (Everitt, Cardinal, Parkinson, & Robbins, 2003). Studies involving human participants also show that lesions of the amygdala disrupt the acquisition of CS-US associations. For example, it has been shown that lesions to the amygdala, but not the OFC, disrupt the development of a conditioned response to a CS predicting white noise (Bechara et al., 1999; Bechara et al., 1995; LaBar, LeDoux, Spencer, & Phelps, 1995). The results involving patients are supported by imaging studies implicating the amygdala in the acquisition of fear conditioning (LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998; Veit et al., 2002).

The amygdala is necessary for some (e.g., passive avoidance learning), but not all (e.g., object discrimination learning and conditional learning), forms of instrumental

learning (Baxter & Murray, 2002). In their recent review, Baxter and Murray (2002) distinguish instrumental learning tasks that require the amygdala from those that do not. They argue that some instrumental learning tasks require the formation of stimulus-*reinforcement* associations and that the amygdala is necessarily involved in the formation of these associations. In contrast, other instrumental learning tasks require the formation of stimulus-*response* associations and that the amygdala is not crucially involved in the formation of these associations. For example, in passive avoidance paradigms, participants learn to respond to rewarding stimuli and avoid responding to stimuli that give rise to punishment. Consequently, passive avoidance learning is thought to rely on the formation of stimulus-reinforcement associations: approach (respond to) stimuli associated with reward and avoid (don't respond to) stimuli associated with punishment (Baxter & Murray, 2002). Lesions of the amygdala disrupt passive avoidance learning (Ambrogi Lorenzini, Bucherelli, Giachetti, Mugnai, & Tassoni, 1991).

In contrast, Baxter and Murray (2002) argue that object discrimination and conditional learning are solved through the formation of a stimulus-response association. Object discrimination learning involves learning to respond to one of two objects (one rewarded and one not rewarded) repeatedly presented in a pair-wise fashion over a series of trials. Thus, the individual can treat the presentation as if it were a compound stimulus: stimulus A/B yields reward if you press B. Conditional learning involves learning to perform a particular motor response in the presence of a particular stimulus (press left button if green light is on, but right button if red light is on). In conditional learning, the individual learns to make response 1 if stimulus A is present and response 2 if stimulus B is present. These two forms of learning differ from stimulus-reinforcement learning because each stimulus is effectively associated with a positive *or* a negative result. Each stimulus is associated with reward provided that the appropriate response is

made in its presence; thus, rather than a stimulus-*reinforcement* association being made to solve the problem, a stimulus-*response* association is sufficient. Lesions of the amygdala do not disrupt object discrimination or conditional learning paradigms (Baxter & Murray, 2002; Burns, Everitt, & Robbins, 1999; Malkova, Gaffan, & Murray, 1997; Petrides, 1985b; 1997).

The amygdala also plays an important role in preferentially allocating attentional resources to emotional stimuli in the environment. Recent formulations suggest that attention is a process involving the competition between multiple stimuli for neural representation (Desimone & Duncan, 1995; Duncan, 1998). Which stimuli win this competition and are "attended to" are a product of both bottom-up sensory processes such as emotional salience and top-down influences such as directed attention (Desimone & Duncan, 1995). Past research has implicated the amygdala in processes involving enhanced attention for emotional information relative to neutral information, even if the emotional information is peripheral to ongoing behaviour (Anderson & Phelps, 2001; Dolan & Vuilleumier, 2003; Vuilleumier, Armony, Driver, & Dolan, 2001). Thus, the amygdala is crucially involved in increasing the perceptual saliency of emotional events, thereby allowing such stimuli to gain preferential access to attentional resources (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Anderson & Phelps, 2001; Vuilleumier et al., 2001).

1.3.4: Dysfunction within the amygdala and psychopathy

The data implicating amygdala dysfunction in psychopathic individuals is considerable (Blair, 2001; Blair et al., 1999; Patrick, 1994). At the anatomical level, structural MRI has revealed an inverse relationship between degree of psychopathic affective and interpersonal traits and amygdala volume (Tiihonen et al., 2000).

Abnormalities reported in functional imaging studies of psychopathy include reduced amygdala activity both during an emotional memory task (Kiehl et al., 2001), and during an aversive conditioning task (Veit et al., 2002). On behavioural measures, individuals with psychopathy show a similar performance pattern as patients with lesions involving the amygdala. Consistent with predictions generated from an amygdala dysfunction explanation of psychopathy, children and adults with the disorder show facial expression recognition deficits (Blair et al., 2001b; Blair et al., 2004b; Kosson, Suchy, Mayer, & Libby, 2002). Evidence suggests that these deficits cross sensory modalities. In a study using auditory emotional expression stimuli, adults with psychopathy demonstrated impaired recognition of fearful vocal affect (Blair et al., 2002).

The abnormalities in cognition extend to conditioning. Like individuals with amygdala lesions, individuals with psychopathy failed to show a conditioned response to stimuli that predict shock (Lykken, 1957). Similar results were obtained by Hare and Quinn, who found reduced skin conductance responding to a tone that had previously predicted shock (Hare & Quinn, 1971). Another study reported that the event related potentials of individuals with psychopathy did not differentiate between a CS+ and a CS- (Flor, Birbaumer, Hermann, Ziegler, & Patrick, 2002), and an imaging study has shown reduced amygdala activity in psychopathic individuals to stimuli predicting pain (Veit et al., 2002). Psychopathic individuals also show reduced augmentation of startle to aversive auditory stimuli preceded by negative images; however, they show appropriate attenuated startle responding when the same auditory stimuli are preceded by positive images (Levenston, Patrick, Bradley, & Lang, 2000; Patrick, Bradley, & Lang, 1993).

Early explorations of instrumental learning in individuals with psychopathy signal impaired performance in at least some forms. In line with suggestions of amygdala dysfunction, individuals with psychopathy show impairment in passive avoidance

learning (Blair et al., 2004a; Newman & Kosson, 1986; Newman & Schmitt, 1998). However, there is a paucity of research concerning other forms of instrumental learning in this patient population.

1.4: Testing the VL/OFC and Amygdala Dysfunction Hypotheses of Psychopathy

1.4.1: Aim of the present thesis

The present thesis was designed to assess the hypothesis that the cognitive impairment evident in patients with psychopathy is attributable to dysfunction within the VL/OFC and the amygdala.

1.4.2: Experiments associated with VL/OFC function

The VL/OFC is involved in reversal learning (Clark, Cools, & Robbins, 2004; Cools et al., 2002; O'Doherty et al., 2001a; Rolls et al., 1994) and decision-making (Bechara et al., 1994; Bechara et al., 1999). In order to examine VL/OFC function, experiments involving response change and decision-making were conducted. Chapter 3 investigates the reversals of stimulus-reinforcement associations on an instrumental learning task. Chapter 4 examines reversals in an object discrimination task, and presents an experiment investigating the performance of individuals with psychopathy on the Iowa Gambling Task. Chapter 5 examines reversals of conditional (stimulus-response) learning.

As discussed earlier in this chapter, similarities have been drawn between individuals with acquired sociopathy and those with developmental psychopathy (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999). In Chapter 6, C.L., a case of acquired sociopathy is presented. Following an injury involving the frontal lobes during

adolescence, CL underwent a dramatic change in personality accompanied by unchecked violence, sexual deviance, and a striking callous emotional style. CL's performance on a battery of standardised and experimental tasks investigating expression recognition, decision-making, instrumental learning, relearning, and social cognition is compared to forensic and community control groups.

In the context of examining instrumental relearning deficits in patients with psychopathy, a more basic question emerged concerning the processes involved in response-reversal. Human studies implicating the OFC in instrumental relearning typically involve abstract secondary reinforcement. Most studies involve situations in which participants must change their responding due to a reversal in reward contingencies associated with available response options. Patterns of OFC activation to these changing reinforcement contingencies are interpreted as evidence of OFC involvement in encoding the reward value of stimuli. However, little is known about the activation of OFC in situations where abrupt changes in the value of reinforcement occur in the absence of demands to shift responding. Chapter 7 presents an event-related fMRI study investigating neural activation to changes in the value of reinforcement together with, and in the absence of, response change. The aim was to address the continuing question of whether the pattern of neural responding to changes in valence is equivalent or dissociable from that associated with response change.

1.4.3: Experiments associated with amygdala function

Due to increasing evidence suggesting that amygdala dysfunction may be a key component of psychopathy, experimental tasks thought to be sensitive to amygdala dysfunction were conducted. Thus, in addition to decision-making and reversal learning, the following chapters will examine the modulation of attention by emotion, stimulus-

reinforcement learning, and expression recognition. The amygdala is thought to play a critical role in the modulation of attention by emotion (Anderson & Phelps, 2001; Dolan & Vuilleumier, 2003; Vuilleumier et al., 2001). As a function of this process, the introduction of an emotional component not relevant to ongoing operant behaviour can have a detrimental impact on performance. An experimental manipulation that utilizes emotional material as distracters should therefore yield superior performance in patients with disorders of emotional processing such as psychopathy. Chapter 2 presents data from such a study.

Recent evidence suggests that different forms of instrumental learning may depend on dissociable neural substrates (Baxter & Murray, 2002). It remains unknown whether individuals with psychopathy show impaired performance on instrumental learning tasks that are not thought to rely on intact amygdala functioning, such as object discrimination and conditional learning. Chapter 3 extends previous suggestions that individuals with psychopathy show impaired stimulus-reinforcement learning in an instrumental learning task with multiple conditioned stimuli. Chapters 5 and 6 address this issue further by including behavioural tasks that contain components that assess the generalisability of the instrumental learning deficit in patients with psychopathy.

Chapter 2

Emotional Responding at the Expense of Cognition

2.1: Introduction

A major focus of neurocognitive investigations involving individuals with psychopathy has been on the emotional basis of the disorder. It has been suggested that a disturbance in the affective processes involved in socialisation can lead to the development of psychopathy (Blair, 1995; Blair, 2003; Frick, 1995; Lykken, 1995). However, although playing a crucial role in many adaptive cognitive functions, emotional responding can also interfere with ongoing behaviour. For example, the introduction of threatening stimuli can result in the cessation of ongoing goal-directed activity. Early animal research has shown that a conditioned stimulus (CS) that predicts shock disrupts adaptive operant behaviour in rats (Estes & Skinner, 1941). Similar processes have been shown in human subjects. For example, the presentation of a CS that previously predicted a burst of white noise resulted in increased response latency during completion of the Tower of Toronto puzzle (Salgado et al., 2000). Vuilleumier and his colleagues demonstrated that the presentation of fearful faces in the periphery whilst participants made same/different judgements on houses resulted in increased response latency even when participants were explicitly instructed to ignore the peripheral face stimuli (Vuilleumier, Armony, Driver, & Dolan, 2001). Finally, a study investigating pictorial stimulus processing revealed slower responding to images with negative relative to neutral content despite the fact that valence was incidental to performing the cognitive task (Simpson et al., 2000).

These data suggest that the introduction of an emotional component not relevant to ongoing behaviour can have a detrimental impact on performance through at least two dissociable processes. One possibility is that emotional material disrupts operant behaviour by eliciting an incompatible response (i.e., freezing/conditioned suppression). Gray has proposed a two-factor learning model that provides a theoretical basis for conditioned freezing/suppression. It postulates that a behavioural activation system (BAS) exists to initiate approach behaviours to appetitive stimuli, and that a behavioural inhibition system (BIS) exists to inhibit behaviour when cues for punishment are in the environment (Gray, 1987). At the neural level, both conditioned freezing and suppression are disrupted following lesions of the lateral and central nuclei of the amygdala in animal studies (LeDoux, Cicchetti, Xagoraris, & Romanski, 1990; Killcross, Robbins, & Everitt, 1997). This process, as conceptualised by Gray, applies to negative, but not positive stimuli, however.

A second possibility is that processing emotional stimuli, regardless of its valence, interferes with goal-relevant stimulus processing. For example, the "integrated competition hypothesis" (Desimone & Duncan, 1995; Duncan, 1998) of attention suggests that competition for neural representation in visually responsive brain systems (sensory and motor as well as cortical and subcortical) occurs when multiple stimuli are present. Both bottom-up sensory processes such as emotional salience and top-down influences such as directed attention determine which stimuli win this competition (Desimone & Duncan, 1995). Past research has implicated the amygdala in processes involving enhanced attention for emotional information relative to neutral information even when the emotional information is peripheral to ongoing behaviour (Anderson & Phelps, 2001; Vuilleumier et al., 2001; Dolan & Vuilleumier, 2003). Imaging studies have reported increased "coupling," or correlated activity between the amygdala and

other regions such as visual cortical areas (Morris, Ohman, & Dolan, 1999), middle occipital areas, fusiform gyri, the superior temporal sulcus, ventromedial, and OFC (Pessoa, McKenna, Gutierrez, & Ungerleider, 2002). This coupling provides support for the idea that a signal originating in the amygdala plays a role in allowing emotional stimuli preferred access to attentional resources. A likely neural mechanism explaining how the emotional control of attention may take place is shown in Figure 2.1, which is derived from the Integrated Emotion Systems model (IES; Blair, 2004).

Figure 2.1: The IES model (Blair, 2004)
This model, derived from portions of the IES model (Blair, 2004), shows a mechanism by which emotional stimuli can disrupt operant behaviour through either the initiation of an incompatible response (i.e., freezing), or the modulation of attention by emotion

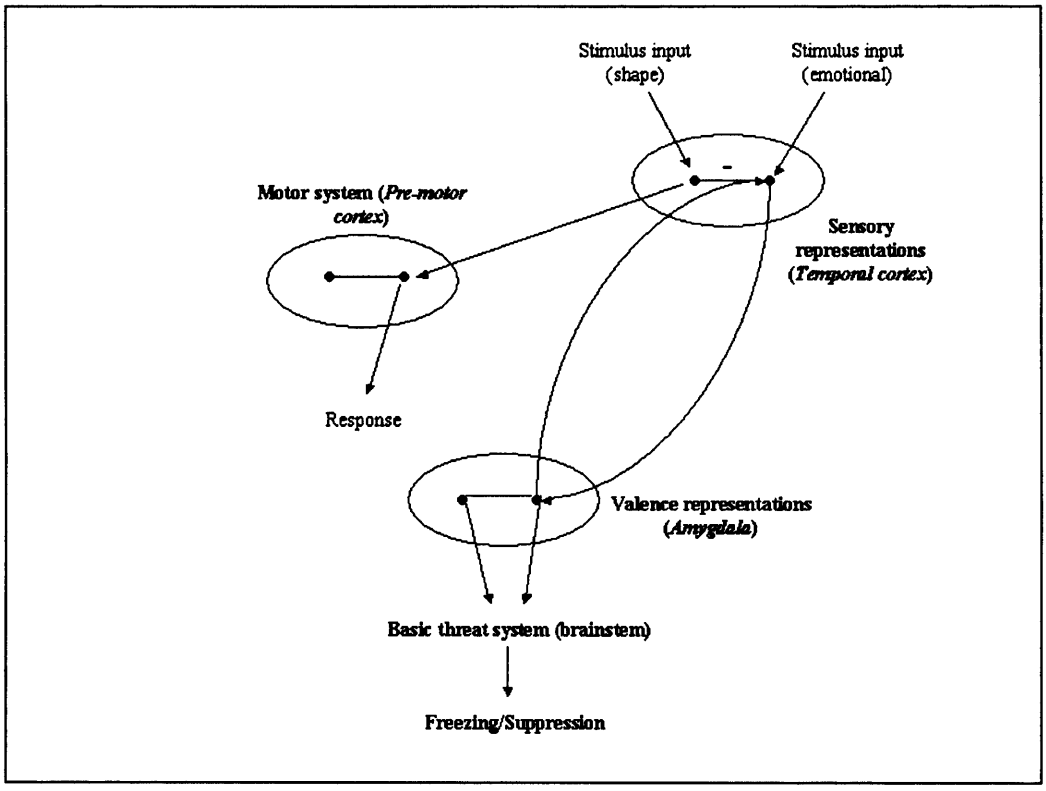


Figure 2.1 depicts circuits involved in generating the "freezing," and in modulating the allocation of attentional resources. In the case of freezing, the lower part of the figure represents a circuit running from medial amygdaloidal areas downward largely via the stria terminalis to the medial hypothalamus and periaqueductal gray. This circuitry modulates the responding to threat as a function of the level of stimulation. At low levels of stimulation, the organism will cease motor activity (Panksepp, 1971). At higher levels, animals will attempt to escape the environment. At higher levels still, where the threat is proximal and escape improbable, the organism will display reactive aggression (Blanchard, Blanchard, & Takahashi, 1978).

The model suggests that emotional stimuli, including conditioned stimuli, activate affect representations in the amygdala that are either innate (for unconditioned stimuli) or formed as a consequence of conditioning. As a result of reciprocal connections between these affect representations and the sensory representations of the emotional stimuli, the sensory activity associated with these emotional representations is augmented. According to Duncan and Desimone's model (1995), representation of one object in the sensory cortex will have an inhibitory effect on the representation of another (thereby creating competition for representation). The strength of a representation is a product of its top-down attentional processes (such as directed attention) or bottom-up processes (such as emotional salience). In essence, when processing capacities are taxed, the representation of one stimulus will occur at the detriment of another. The presentation of multiple stimuli results in reduction in the amount of processing resources afforded to each. As a result, competing stimuli will not receive the same degree of representation, analysis, or control of behaviour (Desimone & Duncan, 1995).

Behavioural studies of psychopathy have focused primarily on situations in which emotional processing deficits result in impaired task performance; however, consistently

impaired performance can be the result of general cognitive factors or reduced motivation rather than a specific neurocognitive deficit. Conditioned suppression and attentional interference paradigms provide an experimental setting in which healthy emotional responding results in slowed or impaired operant responding. This chapter presents a novel experiment, the Emotional Interrupt Task, for which normal emotional responding to positive and negative visual stimuli results in diminished performance on a primary motor response task. The inclusion of positive and negative stimuli enables testing of whether the emotional processing impairment seen in individuals with psychopathy is valence specific. According to the model depicted in Figure 2.1, the modulation of responding can occur through two routes. Projections leading from the amygdala through the brain stem could result in increased response latency by activating the systems involved in conditioned suppression or freezing. If this were to occur, however, one might expect modulation in responding to the aversive stimuli rather than positive stimuli. A second route whereby projections from the amygdala to temporal cortex strengthens the representation of emotional material at the expense of adjacent neutral material is also portrayed in Figure 2.1.

From the perspective of the emotional modulation of attention, both positive and negative stimuli are expected to gain preferential access to attentional resources, particularly as the amygdala has been implicated in the processing of both positive and negative stimuli (Garavan, Pendergrass, Ross, Stein, & Risinger, 2001; Baxter and Murray, 2002; Liberzon, Phan, Decker, & Taylor, 2003; Siebert, Markowitsch, & Bartel, 2003). Interestingly, the data concerning individuals with psychopathy is mixed. Some studies suggest that impairment is restricted to aversive stimuli (Fowles, 1988; Patrick, 1994; Levenston, Patrick, Bradley, & Lang, 2000; Pastor, Molto, Vila, & Lang, 2003) and

others suggest that the impairments extend to appetitive stimuli (Williamson, Harpur, & Hare, 1991; Lorenz & Newman, 2002; Verona, Patrick, Curtin, Bradley, & Lang, 2004).

The objective of the current study was twofold. First, the experiment explores whether the affective modulation of motor responding or attention thought to depend on the amygdala would be reduced in individuals with psychopathy. Second, the task was designed to test whether this modulation would be valence specific. It was predicted that for the comparison group, the introduction of either positive or negative stimuli would result in increased response latencies relative to neutral images. In contrast, it was predicted that for individuals with psychopathy, no modulation of responding would be evident for negative images. Due to contradictory evidence in the literature, the effect that positive images might have on responding in psychopathic individuals was less clear.

2.2: Methods

2.2.1: Participants

Participants were 35 males (16 individuals with psychopathy and 19 comparison individuals) selected from three category B (high security) forensic institutions in the London area. Files were pre-screened to exclude individuals who were older than 55 or whose psychiatric reports revealed a diagnosis for psychosis, organic brain damage, or neurological disorder. All participants were informed that participation was voluntary, and would not affect individual status or record within the institution. Participants did not receive any financial or other gain for their participation or performance on the task. The ages of the participants ranged from 22 to 54 years with a mean of 32.20. The Raven's advanced progressive matrix (Set I) was administered to provide an estimate of intelligence (Raven, 1965; 1995). Raven's scores ranged from 5 to 12 with a mean of 8.17 (according to UK norms, a score of 9 is equivalent to the 50th percentile for

individuals aged 28 to 32 and between 33 and 37; Ravens, 1994). There were no significant group differences in either age or Raven's score. The sample was made up of 28 Caucasian, 1 Pacific Islander, and 6 Afro-Caribbean participants; 2 and 4 Afro-Caribbean participants were in the comparison and psychopathic groups respectively. Participant details by group are presented in Table 2.1.

Table 2.1: Emotional Interrupt Task participant characteristics

	Individuals with psychopathy (n = 16)			**Comparison individuals (n = 19)**		
	M	***SD***	**Range**	***M***	***SD***	**Range**
Participant characteristics						
Ravens Score	8.00	1.55	6 to 10	8.31	2.11	5 to 12
Age	33.44	9.11	22 to 54	31.16	10.02	22 to 54
PCL-R Factor 1 Score	11.66*	1.92	7.50 to 14.90	1.97	1.52	0.00 to 5.00
PCL-R Factor 2 Score	14.90*	0.89	13.05 to 16.00	4.19	3.54	0.55 to 12.10
PCL-R Total	31.87*	1.76	30.00 to 36.80	7.43	4.65	1.65 to 17.60

M = mean; *SD* = standard deviation; * $p < 0.001$.

2.2.2: Measures

2.2.2.1: The PCL-R (Hare, 1991)

The PCL-R consists of 20 behavioural items that are scored on the basis of a file review and semi-structured interview. High inter-rater reliability coefficients for total scores (not less than 0.83) and high Cronbach alpha coefficients and inter-item correlations provide support that the PCL-R is a reliable index of psychopathy in forensic settings (Hare, 1991). The PCL-R consists of two factors identified by psychometric analyses (e.g., Hare, 1991). Factor 1, the affective-interpersonal component, includes items that describe traits central to the classic clinical descriptions of the psychopath;

these include callousness, grandiosity, and a diminished capacity for remorse. Factor 2, the behavioural component, is made up of items that describe traits and behaviours associated with an unstable and antisocial lifestyle; these include impulsivity, poor behavioural controls, and criminal activity. In accordance with the literature and the guidelines of the PCL-R, a diagnosis of psychopathy was given to individuals with a PCL-R score of 30 or above, whereas the comparison individuals had scores of 20 or less. Individuals scoring in the middle range on the PCL-R were excluded due to suggestions that the mixture of psychopathic and non-psychopathic traits that such individuals possess make them inappropriate for either classification (Hare, 1991). After consent was obtained, the participants were interviewed. Participants who declined an interview, but were willing to participate in the experiment, were scored according to file notes (n = 4 split evenly between groups). Evidence suggests that the PCL-R can be scored reliably and validly if detailed file information is available (Wong, 1988; Hare, 1991). Inter-rater reliability was established by means of a Spearman rank correlation conducted on 33 inmates who were scored independently by two raters. The correlation, *rranks* = 0.94 ($p < .001$) is comparable to that presented in the literature (Hare, 1991).

2.2.2.2: The Emotional Interrupt Task

This novel task was designed to assess the impact of emotional processing on relatively habitual motor performance. Participants are engaged in a simple motor response task (left and right button presses to circle/square stimuli). Participants were seated in front of a computer screen and were instructed to respond with their left and right index fingers. Each trial consisted of a fixation point, a pictorial stimulus taken from the International Affective Picture System (IAPS, Lang & Greenwald, 1988), a shape (circle or square stimulus), and the same IAPs image that had preceded the shape.

Pictorial stimuli were classified according to normative ratings provided in the IAPS manual (Lang & Greenwald, 1988); each picture is rated on a 9-point scale on dimensions of valence (higher scores corresponding to greater pleasantness) and arousal (higher scores corresponding to greater arousal). One third of the images were of neutral content (n = 16), one third of the images were positively valenced (n = 16), and one third of the images were negatively valenced (n = 16). The mean valence and arousal ratings for neutral images were 4.76 (ranging from 4.44 to 4.96) and 2.33 (ranging from 1.55 to 2.8) respectively. For positive images, the mean valence and arousal ratings were 7.17 (ranging from 6.58 to 8.14) and 5.56 (ranging from 5.04 and 6.86) respectively. The negative images had a mean valence and arousal rating of 2.63 (ranging from 1.88 to 3.39) and 6.16 (ranging from 5.00 to 7.10) respectively. The negative and the positive stimuli both had significantly higher arousal ratings than the neutral stimuli ($F(1,30) = 310.16$; $p < 0.001$, and $F(1,30) = 412.98$; $p < 0.001$ respectively). Furthermore, the negative stimuli had significantly greater arousal ratings than the positive stimuli ($F(1,30) = 6.31$; $p < 0.05$).

The components of each trial were presented serially; only one component appeared on the screen at any time. The sequence of each trial was: (1) fixation point (800 ms); (2) an IAPS image (200ms); (3) the target stimulus (a circle or square; 150ms); (4) the same IAPs image (400ms); and (5) the inter-trial interval (a blank screen; 1200 ms). Each IAPS image was used once per block. Because there were four blocks, each image appeared twice with each target shape. The total number of trials was therefore 192 (64 neutral, 64 negative, and 64 positive images).

Participants were instructed to respond as quickly as possible to the circle/square stimuli (left button press to circle, right to square). Although participants were told not to respond to the IAPs images, they were asked to attend to them as they might be asked

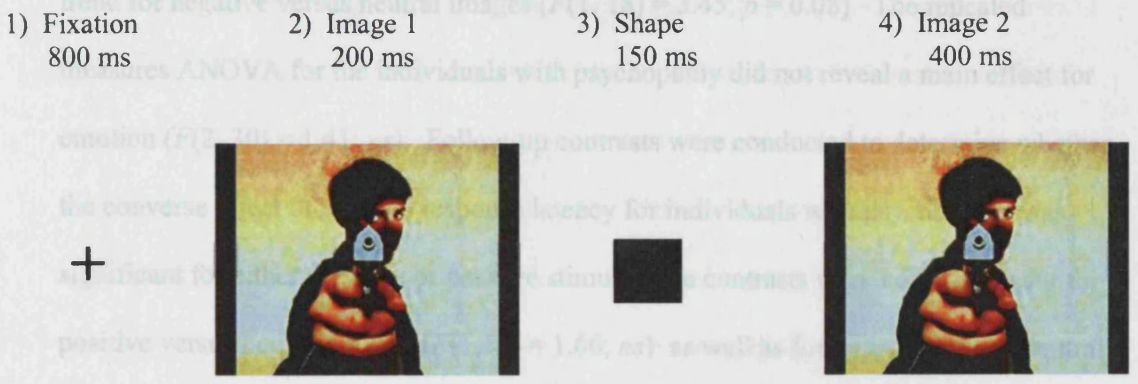
about them later. A sample set of instructions read verbatim to each participant is as follows:

> *In this task, you will be presented with a series of colour pictures depicting people and objects and you will be presented with small black shapes (circles or squares). You will not have to do anything when you see a colour picture of a person or object, but please pay attention to them as you may be asked about them later. When you see the shape, please respond as quickly as you can by pressing the "F" key if you see a square, or press the "J" key if you see a circle. For each series of presentations, you will always be presented with a colour picture, followed by either a small black circle or a small black square, which is followed again by the same colour picture that preceded it. Remember, only respond to the small black circle or square. Speed is very important, so please be sure to press the appropriate key as quickly as you can. You will start with a short practice session to make sure you get the hang of it before the full task begins.*
>
> *The shapes will only be flashed up on the screen very briefly so it is important to be ready for them. Before each trial a cross will appear on the screen to warn you that a trial is about to begin. Please keep your eyes focused on the cross each time and avoid moving them throughout each trial. To help you react as quickly as possible, please keep your index fingers resting on the "F" and "J" keys.*
>
> *So remember, pay attention to each picture, and press the correct key in response to the black and white shapes as fast as you can: F for a Square and J for a Circle.*[1]

The dependent measure was reaction time. Trials for which the reaction time was less than 150 milliseconds or greater than 1500 milliseconds were considered outliers and excluded from the analysis. Figure 2.2 provides a sample schematic of the components of a single trial.

[1] The button that corresponded to the shape was counterbalanced across subjects within each group.

Figure 2.2: Composition of a single trial in the Emotional Interrupt Task



2.3: Results

2.3.1: Response latency data

Table 2.2 shows the mean response latency in milliseconds for each condition for the comparison group and for individuals with psychopathy. An initial 2 (Group: individuals with psychopathy and comparison individuals) X 3 (Valence: Neutral, negative, or positive) repeated measures ANOVA was conducted on the response latency data. There was no main effect for Group ($F(1, 33) = 0.69$; *ns*). The effect for emotion was also non-significant ($F(2, 66) = 0.94$; *ns*); however, a significant Group by Emotion interaction emerged ($F(2, 66) = 3.71$; $p < 0.05$). The comparison group showed a greater response latency when the target stimuli were temporally bracketed by emotional images relative to neutral images; individuals with psychopathy did not show this effect (see Table 2.2). Follow-up within subjects contrasts were conducted with neutral stimuli as the reference level to clarify the nature of the interaction. These revealed significant Group by Emotion interactions for negative versus neutral stimuli ($F(1, 33) = 5.62$; $p< 0.05$) and positive versus neutral stimuli ($F(1, 33) = 6.44$; $p < 0.05$). In order to further explore the nature of these interactions, separate within Groups repeated measures ANOVAs were conducted. For the comparison group, a significant main effect for

emotion was revealed ($F(2, 36) = 3.18$; $p = 0.05$). Follow-up contrasts revealed significant effects for positive versus neutral images ($F(1, 18) = 5.42$; $p < 0.05$), and a trend for negative versus neutral images ($F(1, 18) = 3.45$; $p = 0.08$). The repeated measures ANOVA for the individuals with psychopathy did not reveal a main effect for emotion ($F(2, 30) = 1.41$; *ns*). Follow-up contrasts were conducted to determine whether the converse effect of reduced response latency for individuals with psychopathy was significant for either negative or positive stimuli. The contrasts were non-significant for positive versus neutral images ($F(1,15) = 1.66$; *ns*) as well as for negative versus neutral images ($F(1,15) = 2.29$; *ns*). This shows that although individuals with psychopathy respond slightly faster to negative images relative to neutral images, there are clearly no significant facilitatory effects of emotion for this group.

2.3.2: Error data

A second 2 (Group: individuals with psychopathy and comparison individuals) by 3 (Valence: Negative, neutral, or positive) repeated measures ANOVA was conducted on the error data. No main effect for Group emerged ($F(1, 33) < 1$; *ns*). However, the analysis revealed a main effect for emotion ($F(2, 66) = 3.17$; $p < 0.05$); participants made most errors when the target stimulus was temporally bracketed by negative images. The Group by Emotion interaction approached significance ($F(2, 66) = 2.85$; $p = 0.07$). Whereas the comparison group showed the greatest number of errors for negative items, the psychopathy group showed little difference in error rates between the three conditions. Given the selective effects seen with the response latency data, separate within-subjects repeated measures ANOVAs were conducted for each group. For the comparison group, a significant main effect for emotion was revealed ($F(2,36) = 6.15$; $p < 0.01$). Follow-up contrasts revealed that the comparison individuals made significantly

more errors when the target stimulus was temporally bracketed by negative images rather than neutral images ($F(1, 18) = 8.71$; $p < 0.01$). However, they did not make significantly more errors when the target stimulus was temporally bracketed by positive in contrast to neutral images ($F(1, 18) = 0.36$; *ns*). In addition, significantly more errors were made when the target stimulus was temporally bracketed by negative images relative to positive images ($F(1, 18) = 6.14$; $p < 0.05$). In contrast to the comparison individuals, a within subjects ANOVA involving only individuals with psychopathy, showed no effect for emotion ($F(2,30) < 1$; *ns*). Follow-up contrasts revealed that the individuals with psychopathy did not make significantly more errors whether the target stimulus was temporally bracketed by negative rather than neutral images ($F(1, 15) < 1$; *ns*), positive rather than neutral images ($F(1, 15) = 1.34$; *ns*), or negative in contrast to positive images ($F(1, 15) < 1$; *ns*). Table 2.2 shows the reaction time and error data by group.

Table 2.2: Error and reaction time data

	Individuals with psychopathy (n = 16)		**Comparison individuals (n = 19)**	
	M	***SE***	***M***	***SE***
Reaction Time				
Neutral Trials	409.35	33.00	405.04	26.69
Negative Trials	398.68	21.69	416.51	27.32
Positive Trials	408.36	20.85	423.20	29.27
Correct Responses				
Neutral Trials	57.06	1.52	58.79	0.99
Negative Trials	56.69	1.74	56.84	0.88
Positive Trials	56.25	1.41	58.53	0.98

2.4: Discussion

The purpose the present study was to examine the impact of emotional material on latency of responding in a simple stimulus-response task. It was predicted that to the extent that the current task generates competition for attentional resources, processing of stimuli of either valence would result in increased response latencies for emotionally bracketed stimuli in healthy individuals, but not in individuals with psychopathy. This pattern of results was observed, which supports the idea that attentional resources can be differentially depleted by the presentation of positive or negative visual stimuli and that this depletion is attenuated in individuals with psychopathy.

At least one other study has highlighted circumstances for which a deficit in emotional processing associated with psychopathy translates into superior performance. In it, reduced response latency for individuals with psychopathy on approach trials containing a cue previously associated with punishment was reported (Newman, Wallace, Schmitt, & Arnett, 1997). Punishment associations were established via an initial go/no-go task in which participants learned to press a button to a string of symbols (letters and digits) only if the string did not contain the letter "Q." Responding while the "Q" was present resulted in monetary loss; responding correctly resulted in monetary gain. During the second phase of the task, participants were presented with a stimulus display containing either 3 letters and one number, or four letters in a rectangle and were instructed to respond only if all the symbols were letters. Extra incentives were given for speed. The letter "Q," now an innocuous stimulus like any other letter, was programmed to appear on 50% of the "go" trials. Latency of responding was greater for comparison individuals but not psychopathic individuals when the cue previously associated with punishment was present; however, this result was restricted to comparisons between

psychopathic and non-psychopathic individuals who scored highly on a measure of trait anxiety. In fact, highly anxious individuals with psychopathy showed facilitated responding in the presence of punishment cues. However, this surprising result is not easily interpreted; no *a priori* theory exists to suggest that only individuals with psychopathy with high trait anxiety would show this pattern of responding.

In Figure 2.1, selected components of the Integrated Emotion Systems model are shown to provide a computational rationale for predicting the current results. In line with human neuro-imaging and animal work, two mutually compatible routes by which an emotional stimulus might interfere with on-going behaviour are suggested. First, negative emotional stimuli might activate the basic system mediating the freeze/flight/fight response to threat. In this instance, freezing would interfere with the motor response to the target stimulus (the circle or square). However, this route will not be activated by positive stimuli. In the present study, positive as well as negative stimuli resulted in increased response latencies in the comparison group. The finding that comparison individuals made more errors for negative, but not positively bracketed target stimuli relative to neutral is more likely a result of speed-accuracy tradeoffs than it is a reflection of the activation of the threat circuitry. It may be that the increased errors associated with negative stimuli are attributable to the slightly (though non-significantly) reduced reaction times. This suggests that threat system activation in isolation does not provide a viable account of the data. An additional or alternative system is necessarily involved.

Based on recent models of attention (Desimone & Duncan, 1995; Duncan, 1998), it was suggested that any representation of peripheral emotional stimuli would, to a greater extent than neutral stimuli, interfere with the representation of a more behaviourally relevant target stimulus. Current views on attention characterise this

process as the competition for representation which occurs when multiple stimuli are present (Desimone & Duncan, 1995; Duncan, 1998). The "integrated competition hypothesis" states that given the brain's finite processing capacity for visual stimuli, not all stimuli in the visual field will be processed to the same degree (Duncan, 1998). The degree to which a stimulus is processed is determined by bottom-up sensory processes such as stimulus salience and top-down influences such as directed attention (Desimone & Duncan, 1995). Within this framework, emotional salience is considered to have a bottom-up influence on attention.

At the anatomical level, the amygdala is necessary for the display of conditioned freezing/conditioned suppression (LeDoux et al., 1990; Killcross, Robbins, & Everitt, 1997). It is also involved in the enhanced perceptual saliency of emotional events, which allows such events to gain preferential access to attentional resources (Anderson & Phelps, 2001; Anderson, Christoff, Panitz, De Rosa, Gabrieli, 2003). The suggestion is that emotional material (i.e., conditioned stimuli) activate valence units in the amygdala (due to prior learning) which, because of their reciprocal connections with the representations of the emotional stimuli, further augment the activity of the representations of the emotional stimuli in temporal cortex. Because sensory representations are mutually inhibitory, suppression of the task relevant information results (i.e., the representation of the square/ circle), which leads to increased response latencies.

On the basis of existing evidence for amygdala dysfunction in psychopathy (Patrick, 1994; Blair, Morris, Frith, Perrett, & Dolan, 1999; Blair, 2001; Blair, 2003), it is suggested that the connections between the sensory representations and the valence representations generated by the amygdala are weaker in individuals with psychopathy relative to the comparison group. Based on current work regarding computational

modelling (Blair et al., 2004a; Budhani, 2004), this may caused by reduced activation of the valence units by basic unconditioned stimuli. Reduced activation of the valence units by a stimulus would mean slower Hebbian learning of the connections between these units and the sensory information. Weaker connections between the sensory representations and the valence representations would mean both reduced activation of the basic system mediating the freeze/flight/fight response to threat as well as reduced reciprocal activation of the emotional stimulus competing for representation with the target stimulus. Consequently, there should be less interference by an emotional stimulus on behaviour. This was, of course, seen in the individuals with psychopathy.

The current study examined the detrimental impact of both positive and negative emotional material on task performance. Individuals with psychopathy did not show an effect for emotion whether the visual image was positive or negative. This suggests that individuals with psychopathy present with equivalent processing difficulties for both positive and negative emotional material. In some respects, this is surprising given suggestions that psychopathy might be associated with reward-dominant responding (Fowles, 1988).

The empirical picture, however, is complicated. Individuals with psychopathy show appropriate suppression of the startle reflex following the presentation of positive visual primes, but show reduced augmentation of the startle reflex following the presentation of negative visual primes (Patrick, 1994; Levenston et al., 2000). These results have since been replicated in a Spanish prison population (Pastor, Molto, Vila, & Lang, 2003). This suggests that individuals with psychopathy are unimpaired in processing positive material. Recent affective priming (Peschardt et al., submitted-b) and decision-making studies (Peschardt et al., submitted-a) have suggested that individuals with psychopathy show impaired processing of both positive and negative material, but

that this impairment is particularly severe for negative material. Other studies utilizing emotional material have observed that individuals with psychopathy show the same deficits for processing positive and negative emotional material. For example, Verona and her colleagues report reduced skin conductance responding in individuals with psychopathy to both positive and negative auditory stimuli (Verona et al., 2004). In a lexical decision-making task, comparison individuals were faster to identify positive and negative emotional words than neutral ones, but individuals with psychopathy did not show this emotional advantage (Williamson, Harpur, & Hare, 1991; Lorenz & Newman, 2002). It has been noted that across these studies, individuals with psychopathy identify positive words more quickly than negative words (an advantage ranging from 21ms to 99ms) on average (Peschardt et al., submitted-b).

In short, it is currently unclear which parameters determine the degree to which positive emotional processing is impaired in individuals with psychopathy. One possibility is that the results are influenced by task sensitivity. Tasks such as the Emotional Interrupt or Lexical Decision Tasks may be more sensitive indices of the impairment in processing positive material in individuals with psychopathy. Of course, such a position currently provides no *a priori* reason for why the case would be, nor does it explain why the impairment for negative material is revealed so consistently across studies. A second more interesting possibility is that a modularity in processing exists between tasks. With reference to the model presented in Figure 2.1, tasks showing greater impairment for processing positive material in individuals with psychopathy (the current task, emotional lexical decision-making and affective priming) involve the reciprocal feedback of information from the valence representations to the sensory representations (which, at the anatomical level, involves amygdala-cortical interactions). Tasks showing minimal impairment for processing positive material in individuals with

psychopathy (suppression of the startle reflex by positive visual primes) involve activation of the basic threat system by the valence representations (at the anatomical level, amygdala-brainstem interactions). The pathology associated with psychopathy might particularly affect the reciprocal feedback of information from positive valence representations to the sensory representations. Of course, this explanation is highly speculative. Further investigations into the parameters determining the degree to which the processing of positive emotional material is impaired in individuals with psychopathy, however, may prove to be useful in understanding this disorder.

2.5: Conclusion

In this study, the impact of emotional stimuli on motor responding in individuals with psychopathy and comparison individuals was tested. The comparison group showed increased response latencies to a target stimulus when this stimulus was temporally bracketed by an emotional rather than a neutral visual image; however, individuals with psychopathy did not show a modulation of responding by emotion. This study thus provides further evidence of emotional dysfunction in individuals with psychopathy but does so by demonstrating superior rather than inferior task performance relative to comparison individuals.

Chapter 3

Instrumental Learning and Relearning in Individuals with Psychopathy

3.1: Introduction

In Chapter 1 reviewed evidence suggesting that "instrumental learning" is not a unitary construct, but rather, a term devoted to several neurally dissociable forms learning. In brief, evidence suggests that instrumental learning tasks that are completed through the acquisition of stimulus-*reinforcement* associations necessarily involve the amygdala (Baxter & Murray, 2002). In contrast, the amygdala is not crucial for instrumental learning tasks that require the formation of stimulus-*response* associations. To date, evidence does not suggest that an analogous distinction exists between reversing acquired stimulus-reinforcement associations and reversing stimulus-response associations. This chapter presents an experimental investigation of the acquisition and reversal of stimulus-reinforcement associations acquired during an instrumental learning task in individuals with psychopathy.

A classic example of stimulus-reinforcement learning is demonstrated by the passive avoidance paradigm. In passive avoidance paradigms, participants learn to respond to rewarding stimuli (such as approaching a particular visual stimulus for food) and avoid responding to stimuli that give rise to punishment (do not approach an object associated with electric shock). Consequently, passive avoidance learning is thought to rely on the formation of stimulus-reinforcement associations (Baxter & Murray, 2002); each stimulus presented is associated with either a positive or a negative outcome. Lesions of the amygdala disrupt passive avoidance learning (Ambrogi et al., 1991). In line with suggestions of amygdala dysfunction, individuals with psychopathy show

impaired passive avoidance learning (Blair et al., 2004a; Newman & Kosson, 1986; Newman & Schmitt, 1998).

Passive avoidance learning stands in contrast with the form of instrumental learning demonstrated in object discrimination tasks. During object discriminations, participants select between two objects presented simultaneously; the same two objects are always presented together. One object is associated with reward and the other with non-reward. Thus, participants can solve the problem on each trial with a series of "if-then" statements: if given the choice between object A versus object B, then choose A. The problem can be solved with reference to both stimuli; consequently, stimulus-reinforcement associations need not be formed. The amygdala is not crucial for performing object discriminations (Baxter & Murray, 2002). For example, rhesus monkeys with selective excitotoxic lesions of the amygdala were not impaired selecting between a series of fixed pairs of visual objects for secondary auditory reinforcement (Malkova et al., 1997a).

Although data concerning individuals with psychopathy has consistently indicated the existence of passive avoidance deficits (Blair, et al., 2004; Newman & Kosson, 1986; Newman & Schmitt, 1998), it has yet to be determined whether this deficit extends to include other forms of instrumental learning. Currently, evidence suggests that children with psychopathic tendencies show no difficulties with object discrimination learning (Blair, Colledge, & Mitchell, 2001). In this chapter, an instrumental learning task involving the formation of stimulus-reinforcement associations is presented. Unlike conditional learning tasks which require making a response in the presence of a reference stimulus (all stimuli yield reward provided the appropriate response is made in its presence), the task presented here depends on stimulus-reward associations (each stimulus itself is associated with a positive or negative outcome). In the Tokens Task, the

participant is presented with four stimuli. Each of the four stimuli can be paired with any of the other stimuli, including itself (i.e., there may be two blue tokens on the screen and the participant has to choose one of these two). This is unlike an object discrimination task where the identical pair is presented serially. Two of the stimuli yield reward when selected, and two stimuli yield punishment when selected. This task presumably could be solved through the formation of stimulus-response associations through a complicated series of "if-then" statements. However, given that there are 10 different stimulus combinations, this is not a very efficient solution. A more efficient solution is to associate reward with two stimuli and punishment with the other two stimuli (i.e., to solve the task by the formation of stimulus-reinforcement associations).

In Chapter 1, evidence was reviewed that suggested that the VL/OFC is crucial for adjusting responding to stimuli when reward contingencies change (Rolls, 2000; Rolls, Hornak, Wade, & McGrath, 1994). Specifically, lesions of the OFC disrupt reversing a response that was previously rewarding in favour of a response that was previously punishing (Dias, Robbins, & Roberts, 1996; Swainson et al., 2000). The Tokens Task also contains a component designed to assess instrumental relearning. Following an initial acquisition phase, two reversal phases take place. Data suggests that response suppression or reversal deficits exist in individuals with psychopathy (Blair, Colledge, & Mitchell, 2001; LaPierre, Braun, & Hodgins, 1995). In these studies, however, reversal learning and response suppression were measured using tasks that may require the formation of instrumental associations that differ from those of the Tokens Task. In the LaPierre et al. study (1995), a go/no-go task was used in which participants acquire a prepotent motor response to visual stimuli and are later asked to withhold responding to these previous "go" stimuli. In Blair et al.'s study (2001), reversals occurred after participants had correctly selected one of two fractal images in a series of same-object

discriminations. In each case, the RR being investigated is characterised by a change in stimulus-*response* association, rather than a change in stimulus-*reinforcement* association such as the one that takes place in the Tokens Task.

One other investigation has been conducted investing the performance of individuals with psychopathy on a variant of the Tokens Task (Fine, 2000). In it, Fine (2000) reports severely impaired performance in individuals with psychopathy. As predicted by studies suggesting a role for the amygdala in stimulus-reinforcement acquisitions (Ambrogi et al., 1991), the same study also showed that a patient with a congenital amygdala lesion was severely impaired in learning the reinforcement value of the stimuli. In contrast, two patients with acquired lesions involving the OFC showed intact acquisition during the instrumental learning phase, but were severely impaired on the reversal phase of the task. This provides additional evidence that, although a fractionation of instrumental learning is evident at the neural level, reversal learning appears to rely on the integrity of the OFC whether the association with the stimulus is one involving reinforcement values or a specific response. The objective of the study presented in this chapter was to examine stimulus-reinforcement learning and relearning in an expanded sample of individuals with psychopathy. Furthermore, specific predictions related to the amygdala and OFC explanations of psychopathy were tested. Amygdala dysfunction explanations of psychopathy must predict that stimulus-reinforcement learning as presented in the first phase of the task should be disrupted. In contrast, if only OFC dysfunction is characteristic of the disorder, intact stimulus-reinforcement but disrupted reversal learning should result. The present study tests these predictions in a sample of psychopathic and non-psychopathic inmates, as well as a group of individuals with intermediate levels of psychopathic characteristics to examine whether

their instrumental learning performance shows a categorical or continuous relationship with psychopathy.

3.2: Methods

3.2.1: Participants

The sample was drawn from 7 forensic institutions for men in England. All participants (n = 88) were assessed using the PCL-R. Written consent was obtained from each inmate who agreed to take part, and all were informed that they were free to withdraw from the study at any time. Participants were assigned to high, medium, and low levels of psychopathic characteristics based on their PCL-R scores. The ages of the participants ranged from 21 to 53 years with a mean of 30.77. The National Adult Reading Test (NART; Nelson & Willison, 1991) and Raven's Advanced Matrix (Raven, 1965) were administered to provide indices of general cognitive functioning. Estimated Full-scale IQ scores generated from the NART ranged from 86.50 to 117.21 with a mean of 102.22 (the 50^{th} percentile is 100 for general population; Nelson & Willison, 1991) . Raven's scores ranged from 4 to 12 (maximum score is 12) with a mean of 7.76 (according to UK norms, a score of 9 is equivalent to the 50th percentile for individuals aged 28 to 32 and between 33 and 37; Ravens, 1994). There were no significant group differences for age ($F(2,85) = 1.43$, *ns*), NART ($F(2,85) = 2.10$, *ns*), or Raven's score ($F(2,85) = 0.70$, *ns*). Participant details by group are presented in Table 3.1.

3.2.2: Measures

3.2.2.1: The PCL-R (Hare, 1991).

The PCL-R was used to assign participants to high, medium, and low psychopathy groups. Additional details concerning the psychometric properties and administration of

the PCL-R are presented in section 2.2.2.1 of Chapter 2. After consent was obtained, participants were interviewed. Participants who declined an interview, but were willing to participate in the experiment (n = 8), were scored according to file notes. Inter-rater reliability was established by means of a Spearman rank correlation conducted on 46 inmates who were scored independently by two raters. The correlation, *rranks* = 0.90 ($p < 0.001$), is comparable to that presented in the literature (Hare, 1991). The PCL-R scores of the participants ranged from 1 to 37, with a mean of 20.20 (*SD* = 9.52). Participants were divided into 3 groups based on their PCL-R score: less than 21 (a non-psychopathic group, n = 40), 21-29 (a sub-threshold psychopathy group, *n* = 30), and 30 or above (a psychopathic group, *n* = 18). The mean PCL-R scores by group are given in Table 3.1.

Table 3.1: Tokens Task participant characteristics

	Individuals with psychopathy (n = 18)			**Sub-threshhold psychopathy (n = 30)**			**Comparison individuals (n = 40)**		
	M	***SD***	**Range**	***M***	***SD***	**Range**	***M***	***SD***	**Range**
Participant characteristics									
Ravens Advanced Score	7.59	1.85	5 to 11	8.13	2.11	4 to 12	7.55	2.31	4 to 12
NART Estimated IQ	99.83	9.63	87 to 116	101.08	8.86	87 to 117	104.14	6.83	91 to 117
Age	33.24	7.95	22 to 53	30.77	7.04	21 to 45	29.73	6.91	21 to 48
PCL-R Factor 1 Score	11.89	2.03	9 to 15	9.08	2.43	3 to 13	3.78	2.64	0 to 9
PCL-R Factor 2 Score	15.10	1.43	12 to 17	12.32	2.90	6 to 16	5.86	3.53	0 to 12
PCL-R Total	31.96	2.14	30 to 37	25.04	2.84	21 to 29	11.27	5.52	1 to 20

3.2.2.2: The Tokens Task

This task consists of three key phases and is designed to assess both the ability to form complex stimulus-reinforcement associations and to provide an index of RR difficulties. In this task, participants were told that they can win points by selecting the (coloured) tokens that appear in pairs on the computer screen. Following token selection, a message appeared on the screen informing the participant of the number of points won or lost. The task included four different token colours (red, yellow, green, blue). Two tokens were associated with winning 300 points, and two tokens with losing 300 points. On each trial, two tokens (approximately 2.5 x 2.5cm square) appeared on the screen simultaneously. The participant selected the token of choice using the computer mouse. Following token selection, a message appeared on the screen to inform the participant of the points that had been lost or gained. A total score message at the bottom of the screen was updated before the next trial began. The task consists of nine blocks of ten trials with each block comprising the ten possible token combinations (four same-colour token combinations and six different-colour token combinations). Following an initial familiarization block in which each token combination was presented (10 trials), three main phases of the task took place: (1) an expectation acquisition phase (EA; trials 11 to 40) in which participants learned the reward value of each stimulus; (2) an initial reversal phase (R1; trials 41 to 70) in which one of the previously rewarded tokens became punishing and one of the previously punished tokens became rewarding; and (3) a second reversal phase (R2; trials 71 to 90) in which the remaining two tokens reversed (one became rewarding and the other punishing). Token choices were scored as follows. The trial was scored as "correct" if the subject chose the positive (300 points) over the negative (-300 points) token. The trial was scored as "incorrect" if the subject chose the negative token. If the tokens were of equal value, the trial was not scored. Furthermore,

in Phases 2 and 3, trials were not scored if the subject had not yet had the opportunity to learn the new value of the tokens that had reversed in value. As in the acquisition phase, if a positive (300 points) token was selected over a negative (-300 points) token, the trial was scored as "correct." The trial was scored as "incorrect" if the subject chose the negative token. If the tokens were of equal value, the trial was not scored. Participants were not informed that the task consisted of different phases and were instructed to gain as many points as possible. The instructions for the Tokens Task were as follows:

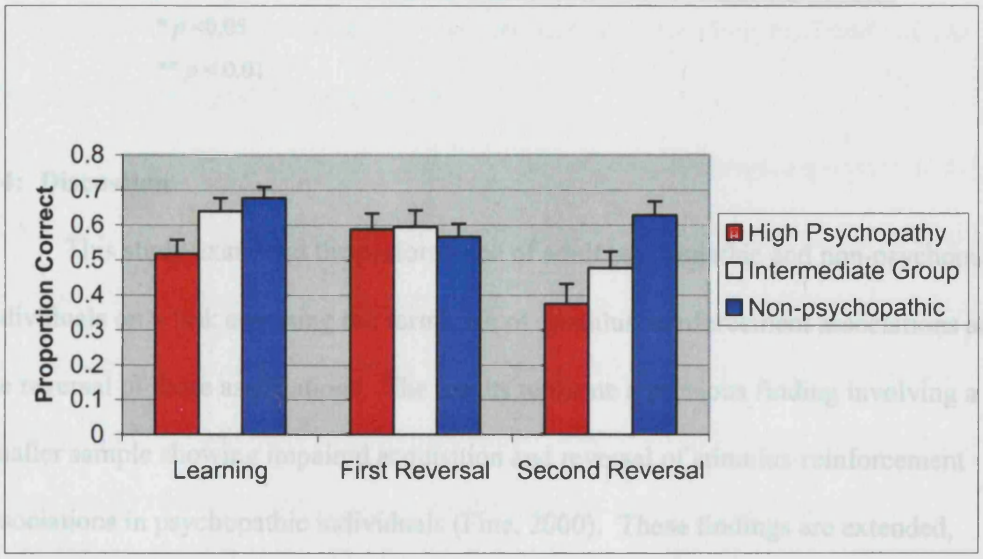
> *Tokens will appear on the screen in pairs. Pressing on some of the tokens will win you points. However, pressing on others will lose you points. When you click on a token you will be told how many points you have won or lost. A total will appear at the bottom of the screen telling you how many points you have won or lost so far. Try to win as many points as possible.*

3.3: Results

Neither age nor estimates of cognitive functioning were correlated with the dependent variables. An initial mixed-model 3(Phase) by 3(Group) Repeated Measures ANOVA was conducted. Mauchly's Test of Sphericity was significant; consequently, the more conservative Greenhouse-Geisser value was used. This revealed a main effect for Phase ($F(1.87,159) = 5.95$; $p < 0.005$), with participants as a whole showing the highest proportion correct at the initial acquisition phase, and the lowest proportion correct at the second RR phase. There was also a significant main effect for group ($F(2, 85) = 5.46$; $p < 0.01$), with individuals with psychopathy making the lowest proportion of correct selections (0.49), the comparison group showing the most (0.62), and the sub-threshold psychopathy group showing an intermediate proportion (0.57). Finally, a Phase by Group interaction was also significant ($F(3.74,159) = 2.87$; $p < 0.05$). While the comparison and sub-threshold psychopathy groups performed best on the initial acquisition phase, the

individuals with psychopathy did not. Furthermore, although the comparison group showed improved performance on the second reversal phase relative to the first, the performance of both the psychopathic and sub-threshold groups declined following the first reversal phase. Figure 3.1 shows mean proportion correct by phase for each group.

Figure 3.1: Proportion correct on the Tokens Task
The bars represent the proportion of correct selections; the vertical lines indicate the standard error of the mean



Follow-up post-hoc tests were conducted to further explore group differences. These tests revealed that individuals with psychopathy performed significantly worse than comparison individuals on the acquisition ($p < 0.05$; *Bonferroni adjusted*) and second reversal ($p < 0.01$) phases. The sub-threshold psychopathy group also made significantly greater errors than the comparison individuals on the second reversal phase ($p = 0.05$).

Correlational analyses were conducted to explore the relationship between psychopathy and performance on the components of the Tokens Task. The results are presented in Table 3.2. As seen in the table, significant negative correlations exist between psychopathy and the proportion correct on the acquisition and second reversal

phase. In contrast, no significant correlations exist for the first reversal phase and either total or factor scores.

Table 3.2: Correlations between PCL-R and number correct

	Factor 1	**Factor 2**	**Total Score**
Acquisition	$r = -0.30^{**}$	$r = -0.28^{**}$	$r = -0.31^{**}$
First Reversal	$r = 0.02$	$r = -0.05$	$r = 0.00$
Second Reversal	$r = -0.26^{*}$	$r = -0.22^{*}$	$r = -0.28^{**}$

$^{*}\ p < 0.05$

$^{**}\ p < 0.01$

3.4: Discussion

This study examined the performance of adult psychopathic and non-psychopathic individuals on a task assessing the formation of stimulus-reinforcement associations and the reversal of those associations. The results replicate a previous finding involving a smaller sample showing impaired acquisition and reversal of stimulus-reinforcement associations in psychopathic individuals (Fine, 2000). These findings are extended, however, by examining the same processes in a group with intermediate levels of psychopathic characteristics. Like psychopathic individuals, this sub-threshold psychopathy group showed significantly greater difficulties relative to the comparison group in reversing their respond when previously rewarding stimuli became punishing and previously punishing stimuli yielded reward. Furthermore, the primary and correlational analyses suggest a continuous rather than discrete relationship between difficulties with both stimulus-reward learning and relearning, and levels of psychopathy. That is, individuals with sub-threshold levels of psychopathy also showed intermediate levels of dysfunction on this form of instrumental learning and relearning.

It is noteworthy that individuals with psychopathy did not show significant reversal deficits in phase 2 (the first reversal phase), but did show significant deficits in the second reversal phase. This finding may be attributable to the failure to learn initial reinforcement associations in the first phase (individuals with psychopathy are performing at chance). The failure to have formed significant reinforcement associations would mean that the first reversal component was no longer an index of response reversal, but instead, effectively an extension of the acquisition phase for the psychopathic individuals. By the third phase, however, individuals with psychopathy had learnt the association to some degree, and consequently, they show profound deficits when required to adjust their responding.

These results complement a growing body of work investigating separable forms of instrumental learning and relearning in individuals with psychopathy. For example, inmates with this disorder show impaired performance in passive avoidance learning (Newman & Kosson, 1986; Newman & Schmitt, 1998), and in extinguishing a once rewarded, but now punished response, (Newman, Patterson, & Kosson, 1987). In the current study, additional evidence is provided that the reversal impairment in psychopathic individuals extends to affect the alteration of stimulus-reinforcement acquisitions.

3.5: Conclusion

As predicted by conceptualizations of psychopathy that implicate an impaired circuit involving the amygdala and OFC, this study showed impaired acquisition of stimulus-reinforcement associations and impaired reversal of these associations, once formed, relative to comparison individuals. In addition, previous findings were extended by the inclusion of a medium scoring group enabling correlational analyses to be

conducted. The significant negative correlations suggest a continuous rather than discrete relationship between psychopathy and both stimulus-reinforcement association acquisition and the reversal of these acquisitions once contingencies change.

Chapter 4

Risky Decisions and Response Reversal:

Is There Evidence of OFC Dysfunction in Psychopathic Individuals?[2]

4.1: Introduction

The experiments described in this chapter are aimed at exploring decision-making and a form of object discrimination learning and relearning described in Chapter 1. Theories of VL/OFC dysfunction in patients with developmental psychopathy were originally inspired by accounts of profound social disturbance shown by some patients with acquired lesions in this area (Anderson, Damasio, Tranel, & Damasio, 2000; Blumer & Benson, 1975; Damasio, 1994; LaPierre et al., 1995). Empirical indications of VL/OFC pathology exist in psychopathic individuals. As previously noted, Lapierre and her colleagues (1995) found that psychopathic individuals made more commissive errors than comparison individuals on a go/no-go task. Furthermore, psychopathic individuals show difficulties extinguishing a previously rewarding response (Newman, Patterson, & Kosson, 1987).

The Iowa Gambling Task is considered a sensitive test of VL/OFC dysfunction (Bechara et al., 1994; Bechara et al., 1999). In this task, participants select cards from four decks each with different rates of monetary reinforcement. Two of the decks contain high rewards, but even higher punishment, making them disadvantageous. The other two decks have low reward, but even lower punishment values. Over the course of the task, healthy participants show a preference for the low-risk decks, whereas patients with

[2] The experiments discussed in this chapter have been published in the article:
Mitchell, D.G.V., Colledge, E., Leonard, A., & Blair R.J.R. (2002). "Risky decisions and response reversal: Is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals?" *Neuropsychologia, 40*: 2013-2022.

lesions to the VL/OFC continue to choose disadvantageously (Bechara et al., 1994; Bechara et al., 1999). A previous study found that psychopathic individuals perform similarly to a comparison group on the gambling task (Schmitt, Brinkley, & Newman, 1999). However, the null result occurred not because the psychopathic group was risk averse, but because the comparison group also failed to show a preference for the low-risk decks. Furthermore, as the authors contend, the studies contained significant procedural differences from the original studies presented by Bechara and his colleagues (1999). Indeed, a recent study that used the recommended procedure showed impaired performance in boys with psychopathic tendencies relative to comparison children (Blair, Colledge, & Mitchell, 2001).

Considerable human and animal data suggest that the VL/OFC is involved in altering behaviour to acquired instrumental associations when they become inappropriate (Dias et al., 1996; Rolls, 2000). Thus, patients with VL/OFC lesions, like psychopathic individuals, have difficulty adjusting their responding when reward contingencies are unexpectedly reversed (Hornak et al., 2004; Rolls et al., 1994). A pure form of reversal learning error is indexed by object reversals. In object reversals, participants choose between two shapes presented to them on a computer screen. Choosing one shape results in reward, whereas choosing the other will result in punishment. Participants, including those with lesions to the VL/OFC, easily learn to choose the correct stimulus. Once the discrimination is learnt, the contingencies reverse so that the previously punished stimulus is rewarded and vice versa. Despite the simplicity of the task, object reversals of this nature have distinguished between patients with lesions to the VL/OFC and those with lesions located in more posterior or dorsal regions of the frontal cortex (Rolls, et al., 1994).

The current study investigates the performance of psychopathic individuals on two tasks thought to be sensitive to VL/OFC dysfunction: the Intradimensional-Extradimensional shift (ID/ED) task and the Iowa Gambling Task. A previous study has suggested that both patients with amygdala lesions and those with damage to the VL/OFC show impaired performance on the gambling task (Bechara et al., 1999). In contrast, the ID/ED task is thought to index two dissociable functions of the frontal cortex. In addition to indexing RR performance, the task also includes a separate extradimensional set (ED) shift component, which requires the participant to attend to a specific aspect of a compound stimulus that had not previously predicted reward or punishment. Whereas impaired RR performance is associated with lesions of the VL/OFC (Dias et al., 1996; Fellows & Farah, 2003; Hornak et al., 2004; Rolls et al., 1994), ED shift dysfunction is associated with damage to the dorsolateral prefrontal cortex (Dias et al., 1996; Owen, Roberts, Polkey, Sahakian, & Robbins, 1991). Moreover, these abilities to perform RRs and ED shifts are doubly dissociable in both animal lesion studies as well as human neuropsychological work (Dias et al., 1996; Fellows & Farah, 2003; Rolls, 2000).

If psychopathy is associated with VL/OFC dysfunction, patients with the disorder should present with impaired performance on both the gambling task and RR components of the ID/ED, but intact performance on the ED shift component. The current study tests these predictions.

4.2: Methods

4.2.1: Participants

Participants were 51 men selected from three Category B (high security) forensic institutions in the London area according to screening and selection criteria described more fully in Chapter 2. Although 51 inmates participated in the study, not all individuals were available for two testing sessions due to inmate transfers. Consequently, 31 adults took part in both tasks, 9 completed only the gambling task and 11 completed only the ID/ED task. The ages of the participants ranged from 21 to 50 with a mean of 33.06 years (SD = 8.03). The mean age for the psychopathic and non-psychopathic groups was 34.42 (SD = 8.07) and 31.64 (SD = 7.91) respectively. The Raven's Advanced Progressive Matrices (Set I) was administered to provide an estimate of intelligence (Raven, 1965; 1995). Raven's scores ranged from 4 to 12 with a mean of 8.12 (SD = 2.47). The mean Raven's Advanced Matrix group scores for psychopathic and control groups was 7.65 (SD = 2.40) and 8.60 (SD = 2.50) respectively. There were no significant group differences in either age ($F(1,49) = 1.55$, *ns*), or Ravens score ($F(1,49) = 1.90$, *ns*). The sample was made up of 49 Caucasian and 2 Afro-Caribbean participants (1 Afro-Caribbean participant in each group). In order to confirm that the groups were age and cognitively matched for the reduced number of participants involved in the ID/ED and Gambling Task primary analyses, additional one-way ANOVAs were conducted on age and Raven's score. These revealed that the mean age and Raven's score did not differ significantly between groups for either task. Table 4.1 lists participant means, standard deviations, and ranges divided by task.

Table 4.1: ID/ED and Gambling Task participant characteristics

Group	PCL-R	Age	Ravens
The Gambling Task			
Psychopathic Inmates	33.10	34.90	7.65
(n = 20)	(2.22; 30 to 37.5)	(7.06; 22 to 47)	(2.48; 4 to 12)
Controls (n = 20)	9.83	30.35	8.65
	(2.88; 4.5 to 14)	(7.41; 21 to 44)	(2.72; 4 to 12)
The ID/ED Task (Primary Analysis)			
Psychopathic Inmates	33.07	32.67	7.73
(n = 15)	(2.29; 30 to 37)	(7.69; 22 to 50)	(2.22; 4 to 12)
Controls (n = 17)	9.16	32.65	8.65
	(3.50; 4.5 to 14)	(7.83; 23 to 45)	(2.67; 4 to 12)
ID/ED Task (All Participants)			
Psychopathic Inmates	32.98	33.62	7.52
(n = 21)	(2.16; 30 to 37.00)	(8.01; 22 to 50)	(2.40; 4 to 12)
Controls (n = 21)	9.08	32.86	8.52
	(3.24; 4.5 to 14)	(7.93; 22 to 45)	(2.44; 4 to 12)

4.2.2: Measures

4.2.2.1: The PCL-R (Hare, 1991)

The PCL-R consists of 20 behavioural items that are scored on the basis of a file review and semi-structured interview as described more fully in section 2.2.2.1. In accordance with the literature and the guidelines of the PCL-R, the psychopathic group for the current study was composed of individuals with a PCL-R score of 30 or above, while the non-psychopathic group was made up of individuals scoring 20 or less.

Participants who declined an interview, but were willing to participate in the experiment ($n = 6$), were scored according to file notes. Inter-rater reliability was established by means of a Spearman rank correlation conducted on 41 inmates who were scored independently by two raters. The correlation, *rranks* = 0.89 ($p < 0.01$), is comparable to that presented in the literature (Hare, 1991).

4.2.2.2: The Iowa Gambling Task (Bechara et al., 1999)

The Iowa Gambling Task was administered in computerised format. The task is a card game in which participants make selections from four decks (labelled A, B, C, and D). Each deck contained different rates and values of monetary reinforcement (play money). Following each selection, the computer emitted a distinct sound (similar to a casino slot machine). Subsequently, a message was displayed on the screen indicating the amount of money the participant had won or lost, and a green bar located above the decks changed proportionately according to the result. The minimum interval between selections was set at 1 second. The computer program terminated automatically after 100 trials (card selections). The participant was not informed in advance about the total number of trials. On the screen, the backs of the cards appeared identical, like real decks of cards. Decks A and C had a higher frequency of punishment, but these punishments were of a lower magnitude. Decks B and D had a lower frequency of punishment, but the punishments were of a higher magnitude. Overall, decks A and B were disadvantageous, and selecting from them would result in a sizable net loss. Ten selections from deck A would yield $1,000, but also five unpredictable losses ranging from $150 to $350 thereby bringing the net loss to $1250. Ten selections from deck B would yield $1000, but also a single loss of $1250. In contrast, ten selections from decks C and D would yield $500, but even smaller losses (ranging from $25 to $75 in deck C and one $250 loss in deck D).

Thus, selecting from decks C and D would result in a net gain of $250. The instructions of the task were read verbatim to each participant as follows:

> *1-In front of you on the screen, there are 4 decks of cards A, B, C, and D.*
> *2-I want you to select one card at a time, by clicking on the card, from any deck you choose.*
> *3-Each time you select a card, the computer will tell you that you won some money. I don't know how much money you will win. You will find out as we go along. Every time you win, the green bar gets bigger.*
> *4-Every so often, however, when you click on a card, the computer tells you that you won some money, but then it says that you lost some money too. I don't know when you will lose, or how much you will lose. You will find out as we go along. Every time you lose, the green bar gets smaller.*
> *5-You are absolutely free to switch from one deck to the other at any time, and as often as you wish.*
> *6-The goal of the game is to win as much money as possible, and if you can't win, avoid losing money as much as possible.*
> *7-You won't know when the game will end. You must keep on playing until the computer stops.*
> *8-I am going to give you this $2000 credit, the green bar, to start the game. The red bar here is a reminder of how much money you borrowed to play the game, and how much money you have to pay back before we see how much you won or lost.*
> *9- It is important to know that the colour of the cards are irrelevant in this game, and there is no way for you to figure out when you lose money. So you must not try to figure out what the computer is doing--you can't!! All I can say is that some decks are worse than the others. You may find all of them bad, but some are worse than the others. No matter how much you find yourself losing, you can still win if you stay away from the worst decks.*

Participants are told that they are free to choose from any of the decks that they wish and that the objective is to win as much money as possible or to avoid losing money as much as possible. In this version of the instructions, it is made clear to the participant that some decks are worse than others and that it is possible to achieve the game objectives if these poor decks are avoided. As with previous studies incorporating the

gambling task, the task was split up into 5 blocks of 20 trials, and the dependent variable was the number of disadvantageous selections made in each block.

4.2.2.3: The ID/ED Task (Dias et al., 1996)

The ID/ED is a multi-component instrumental learning task. Participants learn to select between two stimuli presented to them on a computer screen based on feedback provided on the display (the words "correct" or "incorrect"). The stimuli presented are novel, and involve up to two dimensions (object shape and line contour). The correct stimulus for discrimination is always specified by one dimension or the other (i.e., shape A whether it is paired with line 1 or line 2; or line B regardless of the shape that it is paired with).

On each trial, the two test stimuli appeared randomly in two of four rectangles positioned towards the perimeter of the screen. In order to make a selection, participants used a mouse to position the cursor over the box containing the relevant stimulus. The task consists of nine stages presented in fixed order, and participants were required to learn each discrimination before proceeding past a given stage. In order to meet the learning criterion, participants were required to make eight consecutive correct selections. Participants did not receive any notification when a stage ended and a new rule was presented, nor were they told how many trials or phases made up the experiment. The nine stages and their descriptions were as follows:

(1) Simple discrimination: between two (pink) shapes (Shape 1 and Shape 2). The participant must learn to respond to Shape 1.

(2) Simple reversal: contains the same two stimuli, but the reward contingencies are reversed. Thus, the participant must inhibit responding to Shape 1 and instead respond to Shape 2.

(3) Compound discrimination (separate): a pair of white line patterns adjacent to each shape is introduced (Line 1 and Line 2). However, the contingencies remain unchanged. The participant should maintain responding to Shape 2 whether Shape 2 is

paired with Line 1 or Line 2. The pink shapes and white line are spatially separated to encourage the subject to perceive them as distinct. Pairing is pseudo-random: the same pairings (e.g., Shape1-Line1 and Shape2-Line2) appear in runs of no more than three trials.

(4) Compound discrimination (superimposed): the white lines are superimposed on the pink shapes for this and all subsequent stages so that transfer learning could not be attributed to spatial learning. The contingencies remain the same. The participant should maintain responding to Shape 2.

(5) Compound Reversal: the same stimuli are used but the contingencies are reversed. The participant must reverse their selection of Shape 2 and respond to Shape 1. As before, participants respond to Shape 1 regardless of whether it is paired with Line 1 or Line 2.

(6) Intradimensional shift: new shapes and lines are introduced (Shape 3 and Shape 4 and Line 3 and Line 4). The participant must learn to select Shape 3 whether it is paired with Line 3 or Line 4.

(7) Intradimensional reversal: the contingencies are reversed. The participant must inhibit the selection of Shape 3 and respond to Shape 4.

(8) Extradimensional shift: new shapes and lines are again introduced (Shapes 5 and 6; Lines 5 and 6). However, in this phase of the task the participant must learn that the stimulus feature that predicts reward is the superimposed line rather than the shape. For example, the participant must learn to select Line 5 whether it is paired with Shape 5 or Shape 6.

(9) ED reversal: the contingencies are reversed. The participant must inhibit selecting Line 5 in favour of responding to Line 6.

The dependent variable was the number of errors made before successfully advancing to the next stage (calculated by the computer). In the event that an individual was unable to pass a given stage before making 16 errors on any given stage, the task was terminated.

Each participant was tested individually in a quiet interview room. Both the gambling and ID/ED tasks were presented as part of a larger neuropsychological test battery delivered in random order. The tasks were described without informing the participant of the investigation's specific objectives and expectations. The instructions were read verbatim to each participant as follows:

Can you see the two patterns? One of these patterns is correct, and one of the patterns is wrong. Have a guess at which pattern is correct. If you have made the right choice the computer will show the word "Correct" in green. If you get it wrong the computer will show the word "Wrong" in red. Keep choosing the pattern you think is correct. There is a rule you can follow to make the correct choice each time, but occasionally that rule will change and you have to be prepared for this, but the change will not happen very often.

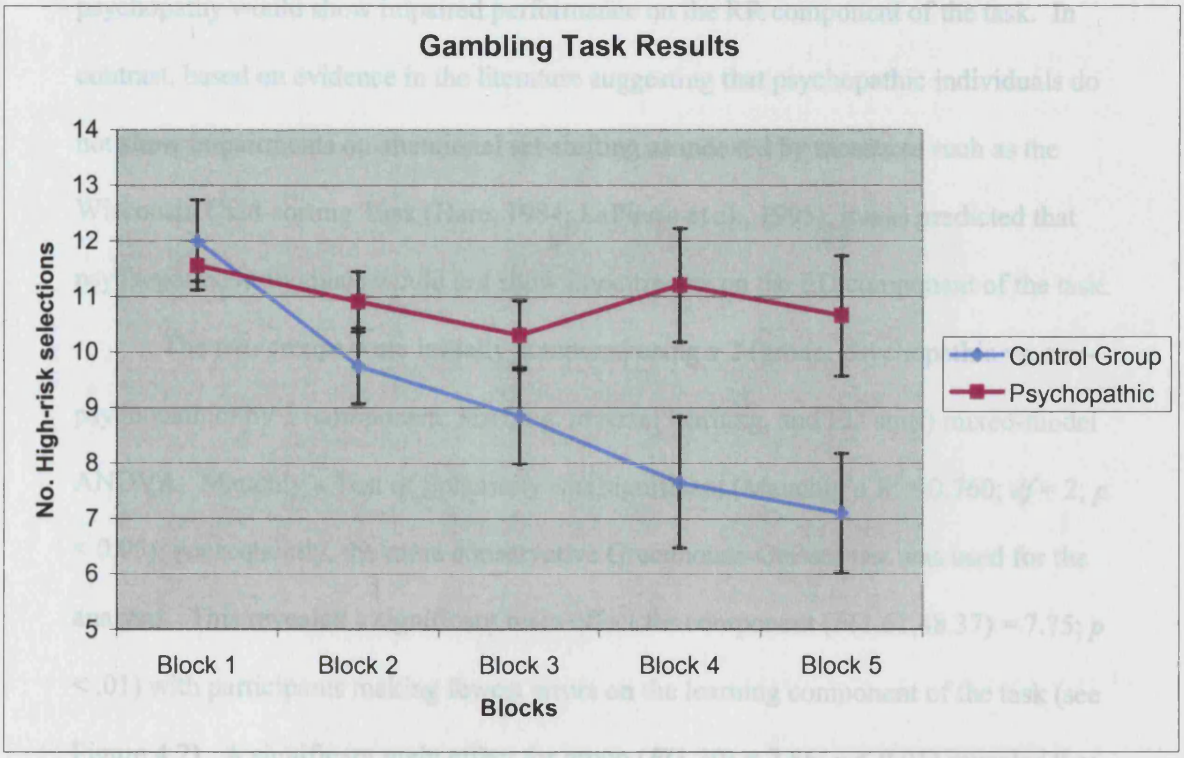
4.3: Results

4.3.1: The Iowa Gambling Task

Based on previous reports in the literature (Bechara et al., 1994; Bechara et al., 1999), it was predicted that the comparison group, but not individuals with psychopathy, would sample randomly at first, but then develop a preference for the advantageous decks. For the analysis, the 100 trials were divided into 5 blocks of 20 trials each. For each block the number of disadvantageous selections (decks A and B) were calculated, as well as the number of advantageous selections (decks C and D). The performance of the two groups is presented in Figure 4.1. The analysis was conducted by way of a mixed model ANOVA. Because Mauchly's Test of Sphericity showed a heterogeneity of covariance, the more conservative Greenhouse-Geisser test was performed. The resulting 2 (control vs. psychopathic) X 5 (blocks of 20 trials) ANOVA yielded main effects for block, (F (2.82, 107.03) = 4.40; $p < 0.01$), with participants becoming increasingly risk-averse over time. There was also an effect for Group, with non-psychopathic individuals making fewer selections from the disadvantageous decks than psychopathic participants ($F(1, 38) = 6.14$; $p < 0.05$). The Block by Group interaction was also significant ($F(2.82, 107.03) = 2.56$; $p < 0.05$; *1-tailed*); non-psychopathic individuals became more risk-averse over time compared with psychopathic participants. An examination of block performance by group in light of the interaction suggests that the main effect for block may best be accounted for by the performance of the control group.

Figure 4.1: Mean number of risky selections on the Iowa Gambling Task

The figure displays the mean number of high risk selections for psychopathic and non-psychopathic individuals by block of 20 trials. Although non-psychopathic individuals learned to avoid the high-risk decks over time, individuals with psychopathy did not show this effect. Vertical lines depict standard errors of the mean.



4.3.2: ID/ED

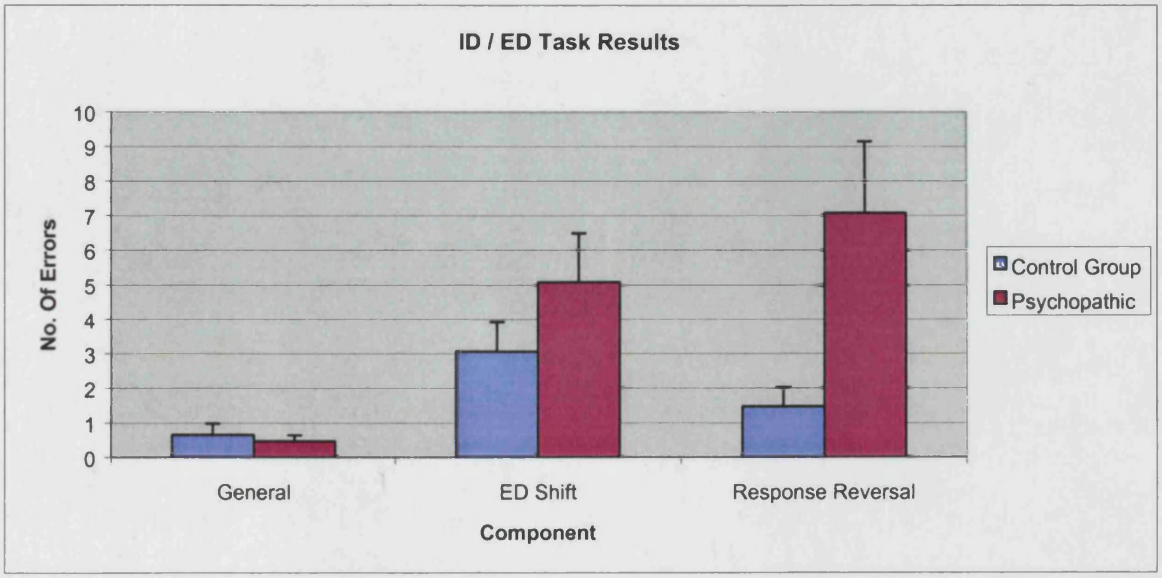
The mean number of errors for each of the three core stage types was calculated for psychopathic and non-psychopathic inmates. The three core stage types were: (1) learning (phases 1 & 6); (2) reversal learning (phases 2, 5, 7, 9); and (3) ED shifting (phase 8). Because obtaining a reliable mean was contingent upon passing all stages, only those participants who had successfully completed each stage were included in the primary analysis. A failure was defined as the commission of 16 errors without achieving the criterion for passing (8 consecutive correct choices). The exception was one individual with psychopathy, who elected to terminate the experiment after making 12

errors on the ED portion of the task. Six psychopathic and four control individuals failed to complete the ED-shift component of the study, which reduced the number of participants included in the primary analysis to 32 (15 psychopathic and 17 non-psychopathic individuals). Based on previous studies investigating RR and extinction (LaPierre et al., 1995; Newman et al., 1987), it was predicted that individuals with psychopathy would show impaired performance on the RR component of the task. In contrast, based on evidence in the literature suggesting that psychopathic individuals do not show impairments on attentional set-shifting as indexed by measures such as the Wisconsin Card-sorting Task (Hare, 1984; LaPierre et al., 1995), it was predicted that psychopathic individuals would not show impairments on the ED component of the task.

The two groups were initially compared using a 2 (group: psychopathic vs. non-psychopathic) by 3 (component: learning, reversal learning, and ED shift) mixed-model ANOVA. Mauchly's Test of Sphericity was significant (Mauchly's $W = 0.760$; $df = 2$; $p < 0.05$); consequently, the more conservative Greenhouse-Geisser test was used for the analysis. This revealed a significant main effect for component ($F(1.61,48.37) = 7.75$; $p < .01$) with participants making fewest errors on the learning component of the task (see Figure 4.2). A significant main effect for group ($F(1,30) = 7.86$; $p < 0.01$) revealed that psychopathic individuals made more errors on the task than the comparison group. A significant group by component interaction also emerged ($F(1.61,48.37) = 3.63$; $p < .05$); psychopathic individuals made greater errors on the reversal component than they did on any other component whereas non-psychopathic individuals made more errors on the ED component than they did on any other component.

Figure 4.2: Mean number of errors on the ID/ED Task

The mean number of errors made by psychopathic and non-psychopathic individuals across key components of the ID/ED Task is presented. Psychopathic individuals made significantly more errors than the comparison group only on the RR component. The bars represent the mean number of errors; the vertical lines indicate the standard error of the mean.



To determine how selective the effect for group was on the components of interest, three one-way ANOVAs were conducted comparing the performance of psychopathic individuals and controls on learning, RRs, and ED shifts. The results showed that psychopathic inmates made significantly more errors on the RR components ($F(1,30) = 7.11; p < 0.05$). In contrast, no significant differences were found for the learning ($F(1,30) = 0.218$; *ns*) or ED components ($F(1,30) = 2.02$; *ns*).

Although no significant difference existed between the two groups with respect to the ED component, ten participants failed to complete this phase and as a consequence, were omitted from the primary analysis. The ED shift performance data was subsequently analyzed in two ways. First, a one-way ANOVA was conducted, which

included participants who had failed the ED component of the task (made 16 errors at which point the task was terminated). This revealed no significant group differences for the number of ED errors ($F(1, 40) = 1.84$; non-significant). Second, in order to assess whether psychopathic individuals were disproportionately represented in this subset of individuals failing the ED component, a chi-square test was conducted. This test revealed that the percentage of psychopathic versus non-psychopathic participants failing the ED component of the task was not significantly different ($X^2(1) = .525$; *ns*).

Table 4.2: Errors on ID/ED by phase

Table 4.2 shows the mean number of errors and standard deviation (in brackets) for psychopathic and non-psychopathic participants for each phase of the ID/ED Task. Significant differences were found only for the first and last reversal phase of the task.

Phase	Psychopathic (n = 21)		Non-Psychopathic (n = 21)	
	Mean Errors (SD)	Number of Failures	Mean Errors (SD)	Number of Failures
1) Simple Discrimination	1.00 (3.26)	0	0.57 (1.36)	0
2) Simple Reversal	3.10* (3.81)	0	0.86 (1.24)	0
3) Compound Discrimination	0.62 (0.86)	0	0.86 (1.59)	0
4) Compound Discrimination Superimposed	0.29 (0.72)	0	0.14 (0.48)	0
5) Compound Reversal	0.29 (0.56)	0	0.29 (0.78)	0
6) ID Shift	0.14 (0.36)	0	0.19 (0.51)	0
7) ID Shift Reversal	0.24 (0.54)	0	0.24 (0.70)	0
8) ED Shift	8.00 (6.66)	6	5.33 (6.07)	4
9) ED Shift Reversal	3.94* (5.96)	0	0.29 (0.47)	0

*$p < 0.05$

4.4: Discussion

The aim of this study was to examine the performance of adult psychopathic individuals on the gambling and ID/ED Tasks. In line with predictions, individuals with psychopathy were less likely to avoid making risky selections over the course of the gambling task relative to comparison individuals. On the ID/ED Task, psychopathic individuals showed a selective RR deficit while performing similarly to control participants on the attentional set-shifting and object discrimination components of the task. This selective result cannot be attributed to a task difficulty effect: a significant interaction revealed that while the psychopathic individuals made more errors on the RR component than the attentional set-shifting component of the task, the opposite was true for the comparison group. Furthermore, both groups performed close to ceiling on the object discrimination (general learning) components of the task. Chapter 3 reviewed evidence that suggests that object discrimination learning does not rely on intact amygdala functioning. Although the simplicity of the object discriminations performed in the ID/ED task limits conclusions that can be drawn about the absence of deficits, the finding provides some additional support for the idea that the instrumental learning deficit evident in individuals with psychopathy is selective.

A previous study has suggested impaired RR performance in psychopathic adults (LaPierre et al., 1995). The present study provides additional evidence for this impairment. To date, one other study has reported intact performance in individuals with psychopathy on the Iowa Gambling Task (Schmitt et al., 1999). However, the procedure of this latter study differed significantly from that used by Bechara and his colleagues (Bechara et al., 1999). In particular, the instructions used by Schmitt and his colleagues did not explicitly state that some decks involved more loss than others and that

participants could win more money overall if they avoided those decks. This may have created a tendency for the comparison group to approach the task as a game of chance, thereby reducing the influence of signals involved in encoding reinforcement contingencies. The only two studies that used the same procedure as Bechara and his colleagues (1999) yielded significant findings. Thus, both boys with psychopathic tendencies in a previous study (Blair et al., 2001a) and adults with psychopathy in the current study show reduced ability to avoid disadvantageous responding.

One way of conceptualising deficits associated with VL/OFC dysfunction is in light of the somatic marker hypothesis (Damasio, 1998; Damasio, 1994). This hypothesis suggests that during decisions of emotional significance, bio-regulatory states provide affective colouring to available response options either through bodily feedback ("body loop") or by representations of bodily feedback formed from previous associations ("as-if body loop"). Effectively, the somatic marker labels an option as either good or bad, thereby rapidly constraining the incentive value of that particular choice. Damasio and his colleagues (Damasio, 1994; Damasio, Tranel, & Damasio, 1991) suggest that dysfunction in the somatic marker system may give rise to psychopathic characteristics. Although the behavioural data in the current study is consistent with the somatic marker hypothesis, the explanation does not predict important findings associated with developmental psychopathy. For example, one of the key findings with respect to the somatic marker hypothesis is that individuals with VL/OFC lesions show reduced autonomic responding to emotionally arousing stimuli (Damasio, Tranel, & Damasio, 1990). In contrast, psychopathic adults and boys with psychopathic tendencies show a selective impairment to stimuli depicting distress, but intact responding to other emotional stimuli such as some threat cues and (sometimes) appetitive stimuli (Blair, 1997; Blair, Jones, Clark, & Smith, 1997; Patrick, Bradley, & Lang, 1993). In light of

these contradictions, it seems unlikely that developmental psychopathy can be explained strictly in terms of somatic marker dysfunction.

It should be noted that the present study has clear significance for another important model of psychopathy--the response modulation hypothesis. Response modulation is defined as a "rapid and relatively automatic (i.e., non-effortful or involuntary) shift of attention from the effortful organisation and implementation of goal-directed behaviour to its evaluation" (Newman, 1998; Newman, Schmitt, & Voss, 1997). Newman (1998) suggests that dysfunction in this system results in a failure to give sufficient consideration to potentially relevant peripheral information when engaging in goal-directed behaviour. Data from the One-Pack Card Playing Task supports the prediction that psychopathic individuals should persist in responding to a previously rewarded response, even if the rate of punishment increases (Fisher & Blair, 1998; Newman et al., 1987; O'Brien & Frick, 1996). However, evidence from lesion studies does not support the idea of a unitary system for response modulation (Dias et al., 1996; Hornak et al., 2004; Rahman, Sahakian, Hodges, Rogers, & Robbins, 1999; Rolls et al., 1994). Collectively, these studies show evidence for a double dissociation between the system crucial for RR and the system crucial for attentional set-shifting. Both RR and attentional set-shifting are clear examples of a modulation in responding. RR requires participants to integrate changes in reinforcement value to redirect their responding from one stimulus to another. The ED shift component of the ID/ED requires individuals to redirect their *attention* from one stimulus dimension (shape) to consider the relevance of another, previously irrelevant stimulus dimension (superimposed design). Thus, the neuropsychological data clearly demonstrate that there is no single system for response set modulation. This conclusion is supported by the current study. The psychopathic individuals in the present study presented with poor RR performance; however, they did

not show deficits in performing ED shifts in the ID/ED, nor is there evidence of attentional set-shifting deficits as indexed by the Wisconsin Card Sorting Task (Hare, 1984; LaPierre et al., 1995). This suggests that the response modulation hypothesis is in need of modification to clarify the parameters under which the system is thought to operate.

An alternative way of conceptualising the current results is in light of the Integrated Emotional Systems (IES) model (Blair, 2004). This model, described in more detail in Chapters 1 and 7, suggests that regions of VLPFC (BA 47) play a crucial role in RR. Functionally, units within BA 47 are thought to detect mismatches between expectations of valence reinforcement (represented by the insula and relayed to medial regions of frontal cortex where they are able to influence behaviour) and actual received reinforcement. Thus, damage to VLPFC, insula, or medial regions of prefrontal cortex will all result in impaired RR. Blair (2004) suggests that the degree of dysfunction in this circuit determines the extent of the RR performance decrement on such measures as the Iowa Gambling and ID/ED Tasks.

It is worth noting that a recent study reports intact RR performance on the ID/ED in boys with psychopathic tendencies (Blair et al., 2001a). Thus, boys with psychopathic tendencies may have less dysfunction than the adults as indicated by their delayed risk avoidance learning but intact performance on the ID/ED. In contrast, the greater dysfunction in psychopathic adults is shown by the lack of risk avoidance learning on the Iowa Gambling Task and the pronounced RR impairment on the ID/ED Task. The differences in the degree of VL/OFC impairment between adults and boys with psychopathic tendencies may be a developmental consequence of the disorder. Given the evidence for interdependence and functional connectivity of the VL/OFC and amygdala, it is possible that a primary deficit within the amygdala could give rise to deficits

associated with VL/OFC impairment. Indeed, a recent imaging study shows reduced amygdalar volume in psychopathic individuals relative to a comparison group (Tiihonen et al., 2000). Although highly speculative, a reduction in afferent input from the amygdala may, over time, have a negative impact on the responsiveness of the VL/OFC. Accordingly, the long-term effects of this dysfunction may not be evident until later in the life-span. This would account for the apparent absence of RR deficits in boys with psychopathic tendencies.

Alternatively, the greater impairment seen in adult psychopathic individuals may arise as a secondary consequence of the behavioural characteristics of the disorder. For example, one of the criteria of psychopathy, stimulation seeking, is often associated with drug use (Hare, 1991). Studies suggest that psychopathy is associated with higher rates of drug abuse, dependence, and poly-drug use (Hemphill, Hart, & Hare, 1994; Smith & Newman, 1990). Furthermore, studies also show impaired performance on the gambling task in alcohol and drug-dependent patients (Bechara et al., 2001; Grant, Contoreggi, & London, 2000). Using a novel decision-making task, Rogers and his colleagues (1999a) assessed the quality of decision-making and deliberation time of individuals with focal VL/OFC damage and individuals who abused amphetamine or opiates. All three groups showed impaired performance on the task relative to comparison groups. Given the neurocognitive impairments associated with chronic drug abuse and the data suggesting that higher rates of abuse and dependence among psychopathic individuals, it cannot be discounted that some of the decision-making impairments seen in psychopathic individuals is acquired as a secondary consequence of their stimulus-seeking behaviour.

4.5: Conclusion

This study provides evidence for impaired performance of psychopathic individuals on the gambling task and the RR component of the ID/ED Task. This impairment may be representative of dysfunction within a neural circuit involving the VL/OFC that is crucial for detecting discrepancies between actual and expected levels of reinforcement. This finding is consistent with other studies that assess psychopathic individuals on tasks with RR or response modulation components (LaPierre et al., 1995; Newman et al., 1987). Interestingly, this dysfunction may be greater in adults with psychopathy than in boys with psychopathic tendencies (Blair et al., 2001a). This disparity raises the possibility that the VL/OFC deficits observed in adult psychopathic individuals may develop as a secondary consequence of early amygdala dysfunction or substance abuse. Further research that controls for substance abuse history and includes decision-making tasks with graded levels of difficulty may establish which of these two possibilities are relevant.

Chapter 5

Psychopathy and Conditional Learning:

Setting the Conditions for the Instrumental Learning Impairment

5.1: Introduction

Chapter 3 described an instrumental learning task believed to rely on stimulus-reinforcement associations. Chapter 4 presented an experiment involving a task that requires participants to demonstrate object discrimination learning. Object discriminations are thought to be a product of stimulus-*response* association formation. The present chapter explores the performance of individuals with psychopathy on an instrumental learning and relearning task—the Ask for Money Task. This paradigm features a third form of instrumental learning referred to as "conditional learning."

Conditional learning is distinct from other forms of learning at the cognitive and neural levels. At the cognitive level, conditional learning involves selecting from available responses in the context of a reference stimulus. In making conditional discriminations, individuals determine their response based on which stimulus is present (i.e., press the left button for stimulus 1, but the right button for stimulus 2). Thus, all stimuli and all responses are associated with reward provided that the appropriate response is performed in the appropriate stimulus' presence. As in object discrimination tasks, conditional learning tasks are not solved through the formation of stimulus-*reinforcement* associations, but rather, through stimulus-*response* associations (Baxter & Murray, 2002). Indeed, conditional discrimination tasks cannot be solved through stimulus-reinforcement learning because each of the stimuli yield reward. Because each

stimulus can yield reward or punishment, a specific affective value or "emotional tone" (i.e., reinforcement value) cannot be associated with the stimulus to facilitate learning.

At the neural level, human and non-human lesion studies support the idea that conditional discriminations represent a dissociable form of instrumental learning. Laboratory rats with basolateral amygdala lesions perform similarly to rats with sham lesions in a conditional learning task; however, the lesioned rats do show impaired second-order conditioning and extinction (Burns, Everitt, & Robbins, 1999). The amygdala is thought to be crucial for representing stimulus value (Everitt, Cardinal, Parkinson, & Robbins, 2003). It is this role in value representation that explains why the amygdala is involved in stimulus-reinforcement, but not stimulus-response learning. Based on these findings, it can be proposed that the amygdala allows affective values to be "tagged" to stimuli to facilitate responding. Affective values are not associated with stimuli in stimulus-response learning; consequently, amygdala function does not facilitate this form of instrumental responding.

Rather than amygdala involvement, the data implicate regions of dorsolateral PFC (Petredes, 1982; 1985a), or inferotemporal cortex together with orbital and ventrolateral structures (Bussey, Wise, & Murray; 2001; 2002) in conditional learning. Macaques with bilateral (but not unilateral) removal of ventral and orbital PFC show conditional learning deficits (Bussey et al., 2001; 2002). Monkeys with crossed disconnection of ventral and OFC with the inferotemporal cortex were also impaired, suggesting that communication between these structures mediates conditional learning (Bussey et al., 2002). It should be noted that the lesions involved in the latter two studies were extensive, including not only ventrolateral regions (Walker's 12), but also regions anterior (Walker's 11), medial (Walker's 11 and 13), and dorsolateral (Walker's 45) to ventrolateral PFC. Monkeys with lesions restricted to the dorsolateral or periarcuate cortex show conditional learning

decrements (Petrides, 1982; Petrides, 1985a). Finally, deficient conditional discrimination performance was observed in humans with lesions involving the frontal cortex (particularly posterior dorsolateral prefrontal cortex), but not in patients with temporal lobectomies (Petrides, 1985b; 1997).

The task described in this chapter consists of two phases. In the first phase, participants encounter one of four stimuli on each trial and learn to choose between two responses for each stimulus. In this initial conditional learning phase, participants learn which response maps on to each of the stimuli. The dorsolateral prefrontal cortex is not thought to be impaired in individuals with psychopathy (Hare, 1984; LaPierre et al., 1995; Blair et al., 2001a; Mitchell et al., 2002); therefore, any focal amygdala or VL/OFC dysfunction analysis of psychopathy must predict intact performance for individuals with psychopathy on the initial phase of this task. In the second phase, response contingencies change for half of the stimuli; participants must alter their responses to these stimuli in order to achieve maximal performance. Regardless of the neural substrate involved in the acquisition of the instrumental association (e.g., the amygdala in the case of stimulus-reinforcement learning and the dorsolateral prefrontal cortex in the case of conditional learning), the VL/OFC is thought to be involved in reversing the acquisition. For example, making object discriminations does not rely on the amygdala, but the VL/OFC *is* nevertheless crucial for the reversal of object discriminations (Rolls et al., 1994; Fellows & Farah, 2003; Hornak et al., 2004). Thus, the VL/OFC is considered to be involved in the second phase of the Ask for Money Task. Consequently, any analysis of psychopathy that emphasizes VL/OFC dysfunction must predict impaired performance for the conditional relearning phase of this task.

In the present study, individuals with psychopathy and a control group are compared on a conditional learning task. Given the lack of evidence for dorsolateral

prefrontal cortex dysfunction in this patient group, it was predicted that individuals with psychopathy would not be impaired in acquiring stimulus-response associations in Phase 1 of the Ask for Money Task. Given the association between VL/OFC dysfunction and psychopathy, however, it was predicted that patients with the disorder would show a selective impairment for conditional relearning.

5.2: Methods

5.2.1: Participants

Participants were 39 individuals selected from three category B (high security) forensic institutions in the London area (19 psychopathic and 20 non-psychopathic offenders). Files were screened to exclude individuals who were older than 55 or whose psychiatric reports revealed a diagnosis for psychosis, organic brain damage or neurological disorder. All participants were informed that participation was voluntary and would not affect individual status or record within the institution. Participants did not receive any financial or other gain for their participation or performance on the task. The ages of the participants ranged from 22 to 54 years old with a mean of 35.54. The Raven's advanced progressive matrix (Set I; Raven, 1965) was administered to provide an estimate of intelligence. Raven's scores ranged from 5 to 12 with a mean of 8.21. There were no significant group differences in either age ($F(1,37) = 0.32$, *ns*) or Raven's score ($F(1,37) = 1.85$, *ns*). The sample was made up of 32 Caucasian and 7 Afro-Caribbean participants; 2 and 5 Afro-Caribbean participants were in the comparison and psychopathic groups respectively. Participant details by group are presented in Table 5.1.

Table 5.1: Ask for Money Task participant characteristics

	Psychopathic Individuals (n = 19)			Comparison Group (n = 20)		
	M	***SD***	**Range**	***M***	***SD***	**Range**
Participant characteristics						
Ravens Advanced Score	7.79	1.60	5 to 11	8.60	2.09	5 to 12
Age	34.63	9.04	22 to 53	36.40	10.40	23 to 54
PCL-R Factor 1 Score	12.76*	1.86	9 to 16	4.10	3.36	0.50 to 12.50
PCL-R Factor 2 Score	14.69*	1.92	7.70 to 16	4.18	2.53	1.10 to 9
PCL-R Total	32.32*	2.23	29.75 to 37.5	9.96	4.51	2.50 to 18.95

*$p < 0.01$.

5.2.2: Measures

5.2.2.1: The PCL-R (Hare, 1991).

The PCL-R was administered as an index of psychopathy. Additional details concerning the psychometric properties and administration of the PCL-R are presented in section 2.2.2.1 of Chapter 2. Inter-rater reliability was established by means of a Spearman rank correlation conducted on 36 inmates who were scored independently by two experienced raters. The correlation, *rranks* = 0.90 ($p < .001$), is comparable to that presented in the literature (Hare, 1991). Individuals with psychopathy were significantly higher on factor 1, 2, and total PCL-R scores than non-psychopathic individuals ($p < 0.01$).

5.2.2.2: The Ask for Money Task (Fine, 2000)

The Ask for Money Task is an instrumental learning task that assesses the ability of participants to adjust responding to changes in reinforcement contingency. In the task, participants were presented with a scenario in which they imagine that they are

attempting to borrow additional money from their relatives. The instructions given to the participants were as follows:

> *You are penniless and desperately need some money. You are going to have to try to persuade your relatives to give you some money. When you see a relative, you can either* ***hint*** *that you would like some money, or* ***beg*** *them to give you some money. Unfortunately, you already owe all of your relatives a lot of money, so sometimes a relative will demand that you repay them some of the money that you owe them.*
>
> *Some relatives may respond better to hinting, and some relatives may respond better to begging. This means that you can maximize the amount of money that your relatives give you, and minimize the amount of money that you pay them back, by learning which is the best way to ask your different relatives for money.*
>
> *In the game, relatives will appear one at a time. You must decide whether to* ***hint*** *for money, or* ***beg****. Once you have chosen, a message will appear on the screen. If your relative has given you money, the message will say* ***"You have won x pounds"****. If your relative has demanded money from you, the message will say,* ***"You have lost x pounds"****. A message will also appear on the screen, telling you how many pounds you would have won or lost if you had chosen the opposite way to ask for money (i.e., hinting instead of begging, or begging instead of hinting).*
>
> *Remember, you need all the money you can get! Good luck.*

The task was presented to the participant via a computer screen. On each trial, a bitmap cartoon representation of a relative appeared in the centre of the screen (see Figure 5.1). Participants could select either the "Hint" or the "Beg" button on each trial by using a mouse. A message would then appear over the button selected indicating how much money the participant had received from their relative, or how much money was reclaimed from them. Additionally, a message would appear on the screen indicating how much money the participant would have gained or lost had they asked for money in the opposite way. Following delivery of reinforcement, there was a four second pause

before the next trial commenced. The total score message was updated on each trial. Four different relatives depicted by distinctive bitmap pictures were in the game. In Phase 1 of the task, two of the relatives were associated with acquiring money and two of the relatives were associated with losing money when the "Beg" or "Hint" buttons were selected.[3] The reward associations are shown in Table 5.2. It can be seen that for two of the relatives, hinting was more advantageous than begging; for the remaining two relatives, the opposite was true. In both phases of the experiment, there were two positive stimuli (relatives who gave money) and two negative stimuli (relatives who demanded an instalment). Each of the phases consisted of blocks consisting of four trials with each relative being presented once in every block in a random order. There were eight blocks (32 trials) in phase 1 (expectation acquisition) and six blocks (24 trials) in phase 2 (reversal). As can be seen in Table 5.2, in the second phase of the task, the optimal strategy for requesting money changed for two of the relatives from begging to hinting for one, and from hinting to begging for the other. The first block of both phase 1 and 2 were not scored and were excluded from the analysis to ensure exposure to the reinforcement contingencies for each stimulus. As a result, phase 1 consisted of 28 scored trials; phase 2 consisted of twenty trials.

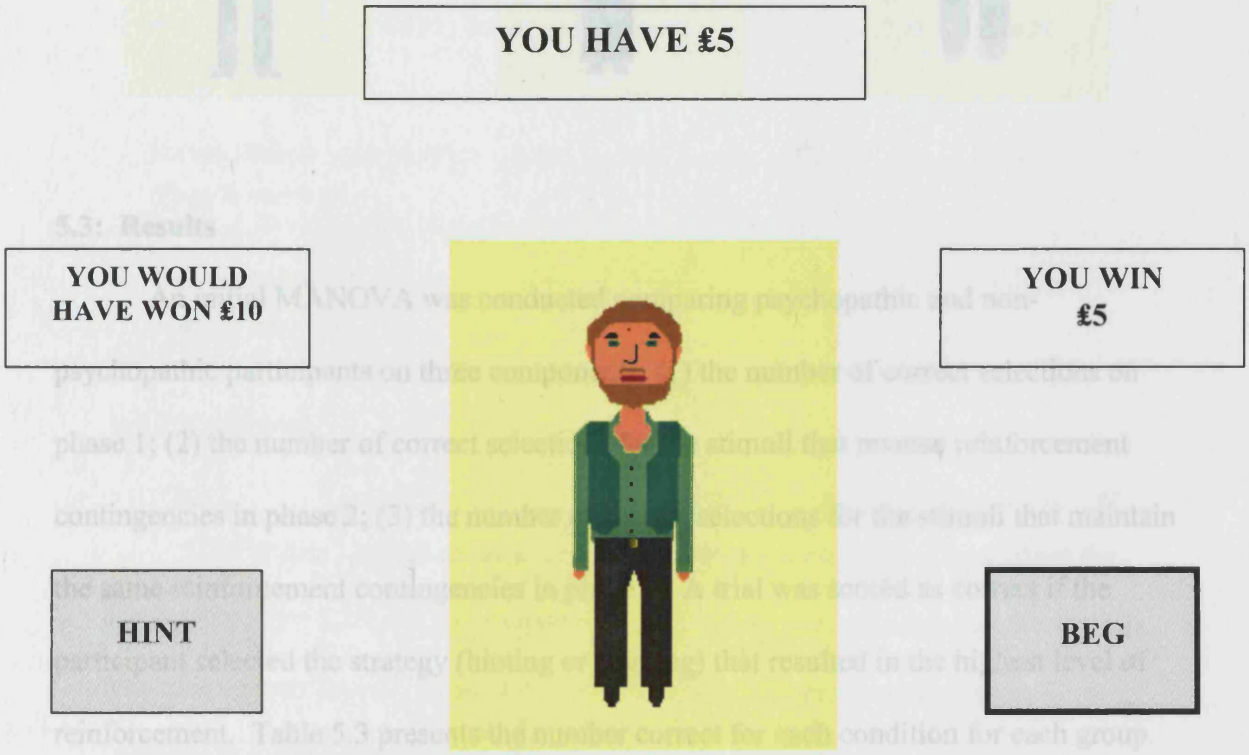
[3] There are differences between the current conditional learning task and tasks used previously in humans by Petrides (1985; 1997). First, in previous conditional learning tasks, stimuli and responses have been paired in a 1:1 matching such that each response is correct for one and only one stimulus. The Ask for Money Task, in contrast, has a 4:2 mapping. Secondly, response choice in this design leads to greater or lesser amounts of reward/punishment for each relative, rather than one response being rewarded and the other punished.

Table 5.2: Response contingencies in each phase of the task

	Phase I		Phase II	
	Hint	**Beg**	**Hint**	**Beg**
A	£5	£10	- £5	- £10
B	- £5	- £10	£5	£10
C	£10	£5	£10	£5
D	- £10	- £5	- £10	- £5

The valence of the response outcomes are reversed between phase one and two for two of the stimuli (positive to negative for relative A and vice versa for relative B). For relative C and D, the valence and values of the response outcomes remain the same. Figure 5.1 depicts the feedback screen of the user interface of the task displayed following a selection that was made.

Figure 5.1: The user interface for the Ask for Money Task (Fine, 2000)
The central figure is the bitmap cartoon depicting one of the conditional stimuli. The response buttons appear on either side of the bitmap image on the lower half of the screen. In this particular example, the participant has erroneously selected the "beg for money" option as depicted by the bolded frame surrounding the "BEG" button.



In the current task, three variables were of particular interest: (1) initial stimulus-response associations as measured by correct selections in phase 1 (conditional learning); (2) behavioural adjustment to changes in reward contingencies as measured by correct selections in phase 2 (for the two stimuli that alter reinforcement values; conditional relearning); and (3) maintenance of old but appropriate stimulus-response associations as measured by correct selections for the two stimuli that retain their original reinforcement contingencies in the second phase.

Figure 5.2: The remaining three conditional stimuli in the Ask for Money Task

5.3: Results

An initial MANOVA was conducted comparing psychopathic and non-psychopathic participants on three components: (1) the number of correct selections on phase 1; (2) the number of correct selections for the stimuli that reverse reinforcement contingencies in phase 2; (3) the number of correct selections for the stimuli that maintain the same reinforcement contingencies in phase 2. A trial was scored as correct if the participant selected the strategy (hinting or begging) that resulted in the highest level of reinforcement. Table 5.3 presents the number correct for each condition for each group.

Table 5.3: Mean number correct by trial type

	Individuals with psychopathy (n = 19)			Comparison individuals (n = 20)		
	M	***SD***	**Range**	***M***	***SD***	**Range**
Correct Selections						
Acquisition (Phase; max = 28)	16.00	4.42	10 to 25	16.90	3.93	8 to 24
Reversed Stimuli (Phase 2; max = 10)	4.95*	1.68	2 to 7	6.15	2.25	0 to 10
Non-reversed Stimuli (Phase 2; max = 10)	5.84	2.06	3 to 10	6.05	2.50	0 to 10

*$p < 0.05$; one-tailed.

The Wilkes Lambda multivariate statistic did not reveal a significant effect for group ($F(3, 35) = 1.20$; *ns*); individuals with psychopathy did not make significantly less correct selections overall. In addition, the individuals with psychopathy and the comparison individuals did not perform significantly different on the conditional learning phase of the task ($F(1,37) < 1$; *ns*). Moreover, one-sample t-tests demonstrated that the performance of both groups was significantly above chance during the conditional learning phase (individuals with psychopathy: ($t(18) = 1.84$; $p < 0.05$, one-tailed); comparison individuals: ($t(19) = 2.96$; $p < .01$); both groups displayed conditional learning.

In phase 2, the individuals with psychopathy made significantly fewer correct selections relative to the comparison individuals for the stimuli that changed reinforcement contingencies ($F(1,37) = 3.54$; $p < 0.05$, $\eta^2 = 0.6$; $d = 0.42$; *one-tailed*).

However, there was no significant group difference for the stimuli that retained the same reinforcement contingencies in phase 2 as they had held in Phase 1 ($F(1,37) < 1$; *ns*). Thus, individuals with psychopathy performed significantly worse than the comparison group at the reversal phase, and did not perform significantly above chance in responding to stimuli for which the response contingency was reversed.

5.4: Discussion

This study examined the performance of adult psychopathic and non-psychopathic individuals on a task assessing conditional learning and relearning. Psychopathic individuals did not show significant impairment relative to a comparison group in conditional learning. However, in line with predictions, the psychopathic group made significantly more errors than comparison individuals on the second phase of the task for the stimuli that reversed their response contingencies. They did not, however, differ from the comparison groups in their responding to stimuli that did not reverse response contingencies.

This experiment joins a body of work investigating different forms of instrumental learning and relearning in individuals with psychopathy. Individuals with psychopathy show difficulty with passive avoidance learning (Newman & Kosson, 1986; Newman & Schmitt, 1998; Blair et al., 2004a), extinction (Newman et al., 1987) and object reversals (Mitchell et al., 2002). Individuals with psychopathy show little difficulty on the current conditional learning task, however, or on reward-only or punishment-only versions of passive avoidance learning (Newman & Kosson, 1986).

Individuals with psychopathy present with impairment in specific aspects of amygdala functioning. Consequently, they should, and do, present with impairment on tasks such as passive avoidance learning, which are reliant on the formation of stimulus-

punishment and stimulus-reward associations. In contrast, individuals with psychopathy do not present with impairment on measures involving stimulus-response associations, such as object discrimination (Mitchell et al., 2002), and the Ask for Money Task.

In light of the present findings, a novel explanation for the dissociation between classic passive avoidance learning and reward or punishment-only variants of passive avoidance learning emerges. Despite consistently showing deficits on traditional forms of passive avoidance learning in which responding to half of the stimuli results in reward and responding to others results in punishment, individuals with psychopathy show intact performance on reward-only and punishment-only versions of the task (Newman, Widom, & Nathan, 1985; Newman & Kosson, 1986). In such versions of the task, participants are rewarded both for responding to the appropriate stimuli and for withholding a response to inappropriate stimuli. When participants respond (or withhold responding) incorrectly, they are not rewarded. As in the conditional learning task described in this chapter, all stimuli can yield reward provided that the appropriate response is made in its presence. Thus, an affective value cannot be tagged to stimuli; the reward-only or punishment-only versions of the passive avoidance are essentially conditional learning tasks and so performance is not facilitated by amygdala responding.

This formulation has already produced testable hypotheses. A recent study involving fMRI found that, in line with predictions, successful passive avoidance learning performance relies on an integrated neural response including the amygdala, hippocampus and medial prefrontal cortex (Budhani, 2004). According to the theory, one must predict that the performance of reward-only/punishment-only versions of the task would activate regions of the dorsolateral prefrontal cortex, but not the amygdala. Preliminary data suggest that this is indeed the case (Budhani, 2004).

Although individuals with psychopathy did not present with significant impairment relative to a comparison group in conditional learning, they did present with difficulties reversing their responding following a contingency change. This was in line with predictions. It is suggested here that VL/OFC is involved in altering both stimulus-reinforcement and stimulus-response associations. The role of the VL/OFC in altering behaviour based on stimulus-reinforcement associations can be seen in extinction paradigms; lesions of OFC in humans and other primates disrupt extinction (Butter, 1969; Rolls et al., 1994). The role of VL/OFC in altering behaviour based on stimulus-response associations can be seen in simple object discrimination RR paradigms; VL/OFC lesions in humans and other primates disrupt performance on such tasks (Rolls et al., 1994; Fellows & Farah, 2003; Berlin, Rolls, & Kischka, 2004; Clark, Cools, & Robbins, 2004). Furthermore, imaging studies investigating reversals of probabilistic object reversals also implicate the VLPFC (Cools et al., 2002; O'Doherty et al., 2003).

5.5: Conclusion

This chapter argued that conditional learning paradigms are reliant on the formation of stimulus-response associations. As predicted by recent conceptualizations of psychopathy that implicate an impaired circuit involving the amygdala and OFC, psychopathy was associated with intact acquisition of conditional discriminations but impaired conditional RRs relative to a comparison group. These results are in line with suggestions that individuals with psychopathy show difficulties with the formation of stimulus-reinforcement but not stimulus-response associations, and in making changes to acquired associations, whether they are between a stimulus and reinforcement or a stimulus and a response.

Chapter 6

Acquired Sociopathy Following an Automobile Accident

6.1: Introduction

This chapter reports on a patient (C.L.), who following trauma to the right frontal region including the OFC at age 14, showed evidence of a personality change consistent with "acquired sociopathy." His behaviour is characterised by impulsivity, promiscuity, pathological lying, temper tantrums, and a callous disregard for the seriousness of his interpersonal behaviour. A series of experimental investigations were conducted to investigate the underlying cognitive dysfunction that may contribute to his aberrant social behaviour. His performance was compared with that of forensic samples of psychopathic and non-psychopathic individuals, and healthy volunteers of comparable age and IQ.

Frontal lobe damage has been associated with profound changes in personality and emotional responding including increased irresponsibility, promiscuity, impulsivity and poor planning (Damasio, 1994). Previously high functioning individuals who incur damage to the OFC are often unable to sustain the same level of occupational or social functioning and often engage in previously uncharacteristic aggression. Chapter 1 of this thesis reviewed evidence suggesting that patients with lesions involving the VL/OFC show deficits on tasks involving: emotional expression recognition (Hornak et al., 1996; Blair & Cipolotti, 2000), empathic responding (Grattan et al., 1994; Eslinger, 1998), and judgements of social convention (Damasio, 1994; Anderson et al., 1999). These laboratory findings are often accompanied by clinical observations of increased emotional volatility, a diminished regard for the welfare of others, and aggression (Burgess & Wood, 1990; Stuss et al., 1992; Damasio, 1994; Anderson et al., 2000).

Patients with OFC damage also show deficient decision-making (Bechara et al., 1994; Damasio, 1994; Bechara et al. 1999) and poor behavioural change following

changes in reinforcement contingencies (Rolls et al., 1994; Rolls, 1997; Rolls, 2004). Currently, three dominant theoretical frameworks exist to characterise key (partially dissociable) functions of VL/OFC. These theories make reference to the VL/OFC as a somatic marker system (Damasio, 1994), a response reversal (RR) system (Rolls et al., 1994; Rolls, 2004), and as a social response reversal (SRR) system (Blair & Cipolotti, 2000). Each of these conceptualisations is reviewed in Chapter 1. In brief, the somatic marker hypothesis suggests that the VL/OFC is involved in integrating feedback from bodily states or representations of bodily states (formed by previous associations) to help precipitate judgements about the value of a particular response option. Rolls (1996) has suggested that the OFC plays a crucial role in modifying behaviour to produce effective reward and punishment-related responding, particularly when reinforcement contingencies change. Thus, the RR explanation suggests that this system is crucial in altering previously acquired stimulus-reward associations when they are no longer appropriate. The SRR explanation suggests that expressions convey important social information used to monitor, assess, and adjust behaviour for ongoing and future interactions (Blair & Cipolotti, 2000). Essentially, this explanation is a social equivalent to the RR position, but with negative social feedback such as angry expressions initiating the cessation of ongoing social behaviour. Individuals who have sustained damage to the OFC are comparably insensitive to these social cues, and so cannot incorporate signals from this system to alter ongoing behaviour.

Recently, a unified model has been proposed that describes how the neurocognitive deficits shown in patients with OFC lesions can lead to reactive aggression (Blair, 2004). In this model, two of the systems described above, the RR and SRR systems, while functionally dissociable, are thought to exist within regions of VLPFC. The RR system is implicated in reactive aggression in its role in detecting

mismatches between the expected and actual value of reinforcement. Frustration, thought to be one of the precursors of reactive aggression (Berkowitz, 1993), results when a behaviour initiated to obtain reinforcement does not result in the expected reward. As a consequence, deficits in the system involved in encoding expectations and violations of these expectations may lead to increased frustration and subsequently, a higher risk for reactive aggression. The SRR is implicated in reactive aggression because of its role in modifying behaviour as a consequence of negative social cues in the environment. This responding is believed to occur as a function of social dominance. The OFC will either mediate or moderate subcortical threat system responding (thereby increasing or decreasing the probability of reactive aggression) depending on how the individuals involved in the interaction are positioned within the dominance hierarchy.

It has been suggested that individuals with early-onset OFC damage present with a syndrome analogous to developmental psychopathy (Anderson et al., 1999). The apparent similarities between the two disorders have prompted individuals to describe patients with lesions to this region as "acquired sociopaths" (Damasio, 1994) or "pseudo-psychopaths" (Blumer & Benson, 1975). In line with this hypothesis, adult psychopathic individuals show deficits on tasks thought to be sensitive to OFC dysfunction. For example, psychopathic individuals showed significantly impaired performance relative to incarcerated controls on a go/no-go task (LaPierre et al., 1995), made significantly more RR errors on the ID/ED Task (Mitchell et al., 2002), and significantly more disadvantageous selections on the Iowa Gambling Task (Blair et al., 2001a; Mitchell et al., 2002). RR deficits in children with psychopathic tendencies appear to be less pronounced, being detected only in tasks with more subtle changes in reward contingencies (Blair et al., 2001a; Budhani, 2004). This finding has prompted speculation

that the RR deficits observed in developmental psychopaths may be a secondary consequence of their disorder (Blair, 2001).

Key differences between developmental psychopathic individuals and acquired sociopaths have been noted. Developmental psychopathy is particularly associated with instrumental *and* reactive aggression (Cornell et al., 1996); however, acquired sociopathy is almost exclusively associated with reactive aggression (Blair & Cipolotti, 2000; Blair, 2004). Evidence suggests that instrumental and reactive aggression are mediated by distinct neurocognitive systems. Reactive aggression, in particular, has been linked to abnormalities within the frontal lobes. Thus, imaging studies have shown reduced prefrontal activity in reactive offenders (Volkow & Tancredi, 1987; Volkow et al., 1995). Raine and his colleagues have shown reduced prefrontal activity in reactive, but not instrumentally violent offenders (Raine et al., 1998).

Recent formulations suggest that early disruption in the socialisation process can be a particular risk factor for instrumental, rather than reactive aggression (Blair, 1995; Blair & Morton, 1995; Blair et al., 1997). Although aversive conditioning does facilitate socialization, it is the distress cues of the victim rather than the threat of punishment that acts as effective aversive stimuli. These aversive stimuli become associated with the aggressive acts that elicit the cues, and through a process of conditioning, the acts themselves become punishing. Given the amygdala's importance in aversive conditioning and instrumental learning (Killcross et al., 1997; LeDoux, 1998; Davis, 2000), dysfunction in this region is suggested to be a specific risk factor for instrumental aggression (Blair, 2001). The link between instrumental aggression and instrumental learning enables the generation of specific predictions with regard to learning deficits in patients with developmental versus acquired psychopathy. Chapters 3, 4, and 5 review evidence suggesting that the amygdala is crucial for some (stimulus-reinforcement), but

not all (stimulus-response) forms of instrumental learning. Consequently, one would predict that the learning impairments demonstrated by individuals with psychopathy would be distinct from those expected in individuals with acquired sociopathy.

A second crucial difference between developmental and acquired psychopathy is that evidence exists for SRR dysfunction in the latter, but not the former patient group (Blair & Cipolotti, 2000). Thus, subjects with developmental psychopathy show intact recognition of some expressions, such as anger (Blair et al., 2001b; 2004b), and appropriate identification of situations likely to induce social disapproval in others (Blair & Cipolotti, 2000). The finding that individuals with psychopathy show intact SRR (Blair & Cipolotti, 2000), but impaired RR (LaPierre et al., 1995; Mitchell et al., 2002) supports suggestions that these systems are dissociable. In fact, a double dissociation has been observed in psychopharmacological research. For example, alcohol and diazepam disrupt SRR (Borrill, Rosen, & Summerfield, 1987; Blair & Curran, 1999), but not RR (Coull et al., 1995). Conversely, serotonergic manipulations modulate RR (Clarke, Dalley, Crofts, Robbins, & Roberts, 2004), but not SRR (Harmer, Bhagwagar, Cowen, & Goodwin, 2001). Despite the presence of RR deficits in developmental psychopathic individuals, the contribution that this dysfunction makes to their instrumental aggression has been questioned (Blair, 2004).

In this chapter, a case study of patient C.L. is presented. Following a severe head injury, C.L. displayed a pattern of behaviour characterised by "acquired sociopathy." The objective of the experiments described in this chapter was to investigate the cognitive impairments that were associated with his injury and help define how these deficits might relate to his profound social and emotional impairment. In light of the debate concerning the neurocognitive impairments associated with developmental psychopathy, the performance of C.L. was compared with that of a small group of individuals with

developmental psychopathy, a comparison group consisting of incarcerated adults, and healthy adults from the community. This enabled comparisons to be drawn between neurocognitive deficits associated with acquired sociopathy and developmental psychopathy. The tasks employed were associated with decision-making, instrumental learning and relearning, emotional expression recognition, and social cognition.

6.2: Case Report

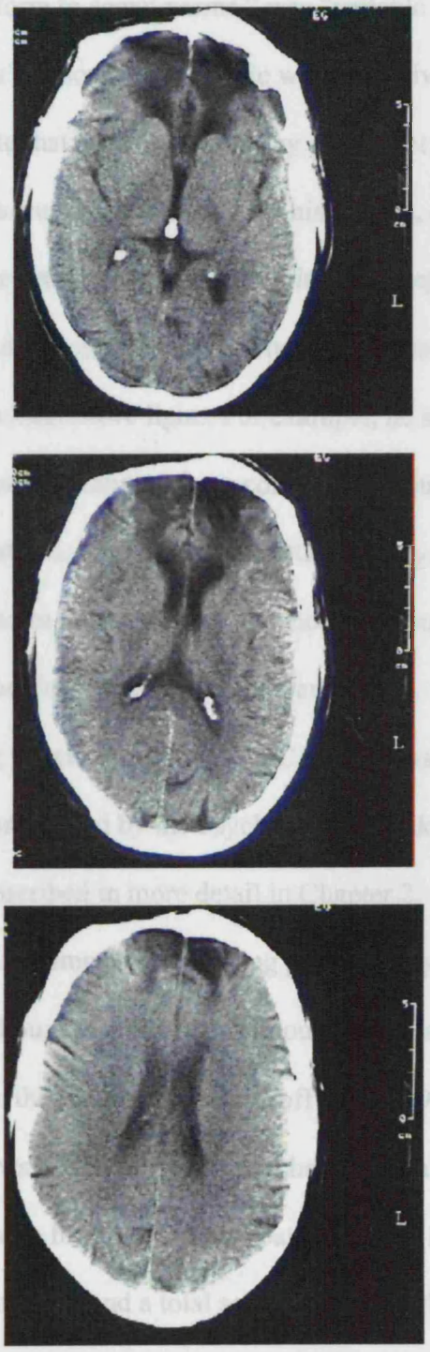
C.L. is a 51 year-old, right-handed male, who was injured at the age of 14 in a traffic accident whilst cycling. He was admitted to hospital suffering severe head trauma in a confused and excited state. Surgeons removed necrotic brain tissue from the right orbital region. A series of operations were undertaken to repair the affected region.

Prior to the accident, C.L. lead a relatively unremarkable childhood. Although engaging in occasional truancy, his records contain no reports of behavioural problems. After his injury, however, he is reported to have lost his friends and to have become socially isolated. He left school without obtaining any qualifications and began to work. His work history thereafter was sporadic, and in his first job he was dismissed in weeks. Subsequently, he obtained roughly a dozen different jobs ranging from one day to six months in duration. His first contact with the law occurred at age 18 when he murdered and sexually assaulted a middle-aged woman. Upon remand, psychiatric evaluations found no evidence of mental illness, but did report the patient's striking failure to appreciate the gravity of his offence. After a few months of observation, the diagnosis "post-traumatic psychopathic disorder" was issued by a psychiatrist, and the patient was described as being a danger both to himself and to others. His recollection of the event was allegedly poor, and he offered several conflicting versions of the offence in an apparent attempt to exculpate himself. His characterisation of the offence ranged from a

full admission to a denial that a murder even took place. During periods in custody, reports indicate that C.L. continued to lack insight into his offence, had an unrealistically high opinion of himself, and that his social interactions were conducted in an abrasive, domineering, and confrontational manner.

C.L. was discharged as an outpatient after being detained for approximately 15 years. In the years that followed, he was in recurrent conflict with the law. He was reported making sexual advances to a minor and, on another occasion, was arrested on suspicion of indecent assault. However, during this period his only conviction was of impaired driving; victims were allegedly reluctant to press charges, and insufficient evidence existed for a conviction. His spouse reported that he abused her both physically and sexually. He was prone to tantrums in which he would throw objects and furniture around the room, particularly after drinking. Individuals responsible for his care felt that their safety was being threatened. Eventually, after several years as an outpatient, C.L. was convicted of a sex-related offence and incarcerated. Specialists reported that his intake period was characterised by persistent lying. Mental health specialists commented on C.L.'s impoverished sense of remorse despite an intact ability to report it verbally. Again, it was suggested that C.L. suffered from psychopathy; however, details about the criteria used to reach this diagnosis were not available.

Figure 6.1: Computerized Axial Tomography (CAT) scan of patient C.L. The scan shows signs of severe necrosis in the frontal cortex. A radiology report concerning a CAT scan of the head reveals wedge-shaped low density lesions involving the frontal lobes bilaterally suggestive of frontal infarcts.



6.3: Behavioural Assessment

In the years following his injury, C.L. qualified for the diagnosis of Antisocial Personality Disorder (ASPD). Reports from the community and within the institution indicate that he "failed to conform to social norms," was "irritable and aggressive," and was "reckless regarding others' personal safety." He was impulsive and sexually inappropriate. Reports indicate that he failed to display any regret or remorse about his prior acts. C.L. also failed to accept responsibility for his actions, often suggesting that he was being victimized by the process. In line with this, C.L. displayed a striking tendency to lie even when his story could easily be checked. These stories usually involved placing himself in an impressive light. For example, he suggested that he had been involved in several professions ranging from complex computer programming to lion-taming and military operations, and while never attaining any formal qualifications, he excelled in each. He seemed to be unaware or unconcerned about the fact that his accounts on one session contradicted accounts he gave earlier.

In addition to fulfilling the criteria for ASPD, C.L. displays striking similarity to the construct of psychopathy as defined by the Psychopathy Checklist-Revised (PCL-R; Hare, 1991). The PCL-R is described in more detail in Chapter 2. It has been reported to be a highly reliable and valid instrument for assessing psychopathy in a prison population (Hare, 1991; Hare, 1996). Although not displaying enough of the symptoms to reach diagnostic criteria according to the recommended cut-off (Hare, 1991), C.L.'s score on the PCL-R reflects that he shows a high degree of psychopathic traits. He attained a score of 13 for Factor 1 (90th percentile for male forensic patients), and 10.1 for Factor 2 (41st percentile for male forensic patients), and a total score of 26.3 (77th percentile). The

score of 26 falls short of a traditional diagnostic cut-off for psychopathy; however, the score suggests significant psychopathic characteristics and places C.L. in the middle-range or "mixed" psychopathy group (Hare, 1991). C.L.'s interpersonal relations are characterised by high rates of poor behavioural controls, sexual promiscuity, grandiosity, superficial relations, and a lack of emotional depth.

6.4: Neuropsychological Assessment

C.L. was assessed using the Wechsler Adult Intelligence Scale-Revised (WAIS-R) and obtained a verbal and performance IQ in the high average range (see Table 6.1). He also obtained a high-average score (9/12) on the Raven Advanced Matrices (Set I; Raven, 1965). C.L.'s performance on the National Adult Reading Test (NART; Nelson & Willison, 1991) was 45/50 signifying an estimated verbal IQ of 124, a performance IQ of 121, and a full-scale IQ of 123.8. On the Recognition Memory Task for words and for faces (Warrington, 1984), C.L. showed intact performance with a score of 46/50 and 33/50 respectively. Performance was near ceiling on the Graded Naming Test (McKenna & Warrington, 1980) and the concrete and abstract word synonym lists with a score of 26/30 in each case (Warrington, McKenna, & Orpwood, 1998). C.L. scored on the 60th percentile on the Rey-Osterriech Complex Figure Test (CFT; Rey & Osterriech, 1993), showing well organised and successful initial drawing order and 70% recall after 40 minutes.

6.4.1: Frontal executive functions

C.L. was assessed on measures of impulse control, planning, and detection of rule change. C.L. made only one error on the Cognitive Estimates Task (Shallice & Evans, 1978), in which participants are asked to give an approximate answer to a general

knowledge question (e.g., "how high is the post office tower?"). He made two minor association errors on the Hayling Task (Burgess & Shallice, 1994) where participants must complete fifteen sentences with an appropriate word, and then complete fifteen sentences with a word that is inappropriate in that context (e.g., "The captain stayed with the sinking" [banana rather than ship]). Five out of the six tasks were attempted on The Six Element Test (Burgess & Shallice, 1993). There were no rule violations, and C.L. performed in the low average range. On the Brixton Test (Burgess & Shallice, 1994), participants learn rules to predict the appearance of a circle in one of 10 locations and also detect rule changes. C.L. made only eight errors, indicating unimpaired performance. On the Wisconsin Card Sorting Task performance all four categories were successfully completed with no more than two errors for each category change showing unimpaired performance. C.L.'s performance on the Proverbs and Classical Weigl Colour Form Task was also unimpaired. Performance on the Stroop and Trail Making Task was at the 17th and 25th percentile respectively, which is indicative of impaired performance given his IQ.

C.L. presented with acquired sociopathy and in the absence of marked frontal executive impairments. In the experiments that follow, the potential neurocognitive basis of C.L.'s acquired sociopathy was explored, and compared with that of developmental psychopathic individuals, forensic controls, and community volunteers. Table 4.1 shows C.L.'s results for the general cognitive and frontal functioning test scores.

Table 6.1: General cognitive and executive function test scores

C.L.	Result
Verbal IQ	115
Performance IQ	113
Full Scale IQ	114
Raven's Advanced Matrix	9/12
Nart	45/50
RM Words	46/50
RM Faces	33/50
GNT	26/30
Concrete Word Synonyms	23/25
Abstract Word Synonyms	23/25
Object Decision	20/20
Cube Analysis	11/11
Cognitive Estimates	Faultless
Wisconsin Card Sorting	Unimpaired
Proverbs	6/8 Pass
Verbal Fluency	60 Pass
Unrestricted nouns	36 (28)
Words beginning "S"	23 (18)
Animals	18 (18)
Weigl Colour Form Sort	Pass
AMIPB, Form I*	
Motor Speed	(40) 25th
Cognitive Speed A	(60) 25th to 50th
Cognitive Speed B	(58) 25th to 50th
Accuracy A	(97%) 50th to 75th
Accuracy B	(100%) > 50th

*AMIPB Information Processing, Form I
Adult average score in brackets

6.5: Experimental Investigation

The following experiments were conducted over a period of 1 year. The tasks were introduced to C.L. as way of understanding his behaviour and determining how brain function is correlated with decision-making and emotional processing skills. He found the task to be a stimulating diversion from normal ward activities. In order to investigate C.L.'s acquired sociopathy, instrumental learning and relearning, emotional expression recognition, and social cognition tasks were carried out.

6.5.1: Comparison groups

C.L.'s performance on social cognition, reversal learning, and emotional expression processing was compared with that of five male prison inmates with developmental psychopathy and five non-psychopathic control inmates held in equivalent forensic institutions. In addition, five healthy males from the community also participated in the study. Psychopathy was assessed in the forensic samples using the PCL-R (Hare, 1991). Files were pre-screened to exclude individuals who were older than 60 or whose psychiatric reports revealed a diagnosis for psychosis, organic brain damage, or neurological disorder. All participants were informed that participation was voluntary and would not affect individual status or record within the institution. Participants did not receive any financial or other gain for their participation. The ages of the forensic participants ranged from 37 to 53 years with a mean of 42.5. The Raven's Advanced Progressive Matrix (Set I) was administered to provide an estimate of intelligence. Raven's scores ranged from 5 to 11 with a mean of 7.70. There were no significant group differences in either age ($F(1,8) = 1.11$, *ns*) or Raven's score ($F(1,8) = 0.68$). Because of the unique demographic profile of the forensic institutions from which the sample was drawn, it was not possible to select a forensic comparison group of C.L.'s age and IQ. Consequently, a second comparison group from the community was gathered. For the healthy community-based volunteers, the age range was 44 to 59 (mean = 50.4). The range of scores on the Raven's advanced matrix was 7 to 12 (mean = 9.60). Participant details are shown in Table 6.2.

Table 6.2: Characteristics of the comparison groups

	Individuals with psychopathy (n = 5)			Comparison individuals (n = 5)			Community Controls (n = 5)		
	M	***SD***	**Range**	***M***	***SD***	**Range**	***M***	***SD***	**Range**
Participant characteristics									
Ravens Score	7.80	1.92	5 to 10	7.60	2.70	5 to 11	9.60	2.07	7 to 12
Age	41.40	6.66	37 to 53	43.60	3.36	39 to 48	50.40	6.88	44 to 59
PCL-R Factor 1 Score	11.07	2.17	9.00 to 14.00	4.40	3.90	1.50 to 10.00	-	-	-
PCL-R Factor 2 Score	15.00	1.87	12.00 to 17.00	5.59	3.59	1.30 to 10.15	-	-	-
PCL-R Total	31.07	1.37	30.00 to 33.35	11.92	7.28	3.30 to 20.00	-	-	-

6.5.2: Decision-making, instrumental learning, and relearning

The tasks in this section are designed to assess behavioural adjustment based on differences or changes in reinforcement. Two tasks investigated separable forms of instrumental learning: the Passive Avoidance Task (Blair, et al., 2004) and the Ask for Money Task (Fine, 2000). Three tasks were administered in order to investigate sensitivity to rates or changes in reinforcement value: the Iowa Gambling Task (Bechara et al., 1994; Bechara et al., 1999) the ID/ED Task (Dias et al., 1996), and the One-Pack Card-Playing Task (Newman et al., 1987).

Task 6.1: The Passive Avoidance Task (Blair et al., 2004a). A classic measure of the ability to learn from aversive experiences is passive avoidance. In this paradigm, individuals learn to respond to stimuli that yield reward and to avoid responding to stimuli that yield punishment. Success on the task is measured by the rates of passive avoidance errors (responses to punishing stimuli) and errors of omission (failing to respond to rewarded stimuli).

In order to perform the task, participants associate a specific stimulus to a specific level of reinforcement; that is, some stimuli are associated with a positive outcome and initiate approach behaviours, and others are associated with a negative outcome and are avoided. Stimulus-reinforcement associations represent a form of instrumental learning for which the amygdala is considered crucial (Baxter & Murray, 2002). Individuals with psychopathy consistently show deficits on classic measures of instrumental learning such as passive avoidance (Newman & Kosson, 1986; Newman & Schmitt, 1998; Blair et al., 2004a). Instrumental learning deficits have not, however, been associated with reactive aggression or OFC pathology (Blair, 2004).

Procedure: The Graded Passive Avoidance Task (Blair et al., 2004a) is a modified version of the Newman and Kosson (1986) task in which individual stimuli (in this case numbers) are associated with specific levels of point reward/punishment. In the task, 8 different numbers were presented to the participant once per block, for 10 blocks. Four of the stimuli yielded reward when selected, and four yielded punishment. The independent variable was the number of passive avoidance errors.

Results: C.L.'s made significantly less passive avoidance errors than the psychopathic individuals (number of errors is less than the 95% confidence interval), and performed within the range of both comparison groups. He also made a similar amount of omission errors as any group, suggesting that his reduced passive avoidance errors was not due to a bias towards non-responding. In line with predictions, no evidence for stimulus-reinforcement learning deficits was present.

Table 6.3: Passive Avoidance Task results

	Passive Avoidance errors	**Omission errors**
CL	8	7
Psychopathic		
M	24.40	11.00
SD	9.45	5.24
Range	10 to 36	5 to 19
CI[4]	12.7 to 36.1	4.5 to 17.5
Forensic controls		
M	18.00	13.00
SD	8.09	7.45
Range	7 to 28	4 to 24
CI	8.0 to 28.0	3.7 to 22.3
Controls		
M	9.40	10.80
SD	4.28	7.82
Range	5 to 16	0 to 20
CI	4.1 to 14.7	1.1 to 20.5

[4] CI = confidence interval. All confidence intervals are calculated at 0.95 level.

Task 6.2: The Ask for Money Task (Fine, 2000). In conditional learning, participants learn to make one of two responses when they encounter each stimulus. This form of instrumental learning is said to show stimulus-response learning rather than stimulus reinforcement learning (Baxter & Murray, 2002). Unlike in stimulus-reinforcement association formation, the reinforcement value of the stimulus for stimulus-response association formation is contingent upon the response. Because stimuli are not intrinsically rewarding or punishing, there is no reinforcement contingency to be learnt. Lesion studies suggest that the amygdala is not crucial for the formation of stimulus-response associations, which are instead disrupted by lesions to dorsolateral areas of prefrontal cortex (Petrides, 1982; 1985a; 1990).

Procedure: The Ask for Money Task is described more fully in Chapter 5. In brief, participants are presented with a scenario in which they imagine that they are attempting to borrow additional money from various relatives. There were four relatives (conditional stimuli represented by unique bitmap images and presented on a computer screen), and two response options (buttons) for each stimulus (hint or beg). Two of the stimuli yielded a favourable response to hinting, and two of the stimuli yielded a more favourable response to begging. After each selection, the computer displayed the amount of money won or lost, but also the amount that would have been gained or lost had the opposite response been selected. The task consists of two phases. In the first phase, participants learn the stimulus-response associations. The second phase measures conditional reversals; half of the stimuli reverse so that the previously disadvantageous response is now advantageous. There were eight blocks (32 trials) in phase 1 (expectation acquisition) and six blocks (24 trials) in phase 2 (reversal). Participants were asked to acquire as much money as possible.

Results: C.L. showed impaired conditional learning. His performance, although within the range and confidence intervals of the forensic samples, falls outside of the range and CI of the age and IQ-matched comparison group. His impoverished performance is not unexpected given the extent of his frontal lesion. The results provide further support for the notion that dissociable forms of instrumental learning exist.

Table 6.4: Ask for Money Task results

	Conditional Acquisition Correct	**Phase 2 Reversed Stimuli**	**Phase 2 Non-reversed Stimuli**
CL	13	4	3
Psychopathic			
M	18.20	4.60	6.00
SD	4.32	1.14	1.87
Range	13 to 24	3 to 6	4 to 9
CI	12.8 to 23.6	3.2 to 6.0	3.7 to 8.3
Forensic controls			
M	15.80	5.80	4.40
SD	4.44	0.84	2.51
Range	8 to 19	5 to 7	0 to 6
CI	10.3 to 21.3	4.8 to 6.8	1.3 to 7.5
Controls			
M	23.00	5.80	9.60
SD	5.61	2.86	1.67
Range	15 to 30	1 to 8	7 to 11
CI	16.0 to 30.0	2.2 to 9.4	7.5 to 11.7

Task 6.3: The Iowa Gambling Task (Bechara et al., 1994; Bechara et al., 1999). The task is described in detail in Chapter 5. Bechara and his colleagues have reported that both patients with acquired sociopathy and individuals with amygdala lesions show impaired performance on the gambling task (Bechara et al., 1994; Bechara et al., 1999). Furthermore, recent investigations have suggested deficits in incarcerated psychopaths (Mitchell et al., 2002), and in a sample of sub-clinical psychopaths (Van Honk, Hermans, Putman, Montagne, & Schutter, 2002). Given the tasks sensitivity to both amygdala and

OFC dysfunction, it was predicted that both individuals with psychopathy and C.L. would show impaired performance relative to the comparison groups.

Procedure: The gambling task was administered in computerised format with a schedule of reinforcement described by Bechara and his colleagues (1999). The task parameters and procedure is described in detail in Chapter 5. In brief, participants make selections from four decks of cards containing different rates of monetary reinforcement. Two of the decks (high-risk decks) involve a high magnitude of reward, but even higher magnitude of punishment resulting in a net loss. The remaining two decks involve a lower magnitude of reward but punishments are of a lower magnitude. Participants must learn to avoid the high-risk decks in order to achieve the game objectives. In this version of the instructions, it is explicit that some decks are worse than others and that it is possible to achieve the game objectives if these poor decks are avoided.

Results: Over the course of the task, CL failed to show avoidance of risky selections. His performance is similar to individuals with psychopathy (mean risky selections 54 versus 52 for the psychopathic group). In total, CL selects more cards from the disadvantageous decks than the advantageous decks. His total number of risky selections is outside the range and CI of the forensic and community control groups. Table 6.5 shows the mean number of risky selections by block for CL, psychopathic individuals and the comparison groups.

Table 6.5: Iowa Gambling Task results

	CL	Individuals with psychopathy				Forensic Controls				Community Controls			
		M	*SD*	Range	*CI*	*M*	*SD*	Range	*CI*	*M*	*SD*	Range	*CI*
Total Disadvantageous Selections													
Block 1	15	10.60	1.34	9 to 12	8.9 to 12.3	13.60	4.45	8 to 19	8.1 to 19.1	9.00	8.31	0 to 16	0 to 19.3
Block 2	8	8.80	2.17	6 to 11	6.1 to 11.5	7.60	4.39	2 to 13	2.1 to 13.1	5.40	3.65	0 to 10	0.9 to 9.9
Block 3	7	10.40	3.71	5 to 13	5.8 to 15.0	6.00	4.53	0 to 11	0.4 to 11.6	6.20	3.42	1 to 10	2.0 to 10.4
Block 4	15	12.40	4.83	8 to 18	6.4 to 18.4	2.00	3.94	0 to 9	0 to 6.9	8.00	6.63	2 to 18	0 to 16.2
Block 5	9	9.80	5.93	6 to 20	2.4 to 17.2	3.00	4.24	0 to 9	0 to 8.3	4.80	3.03	3 to 10	1.0 to 8.6
Totals by Decks													
Decks A/B	54	52.00	12.49	39 to 67	36.5 to 67.5	32.20	11.58	20 to 50	17.8 to 46.6	33.40	10.48	22 to 46	20.4 to 46.4
Decks C/D	46	48.00	12.49	33 to 61	32.5 to 63.5	67.80	11.58	50 to 80	53.4 to 82.2	66.60	10.48	54 to 78	53.6 to 79.6

Task 6.4: The Intradimensional-Extradimensional Shift (ID/ED) Task (Dias et al., 1996). The ID/ED Task measures four separable cognitive skills: (1) the ability to perform simple and compound discrimination learning; (2) the ability to transfer this learning to a new exemplar of the same dimension (Intradimensional shift); (3) the ability to change response set criteria from one dimensional to a second previously irrelevant dimension (Extradimensional shift); and (4) the ability to alter responding when the reinforcement contingencies reverse (RR).

Studies have shown that patients with OFC lesions exhibit impaired performance on the RR component while dorsolateral lesions are associated with ED shift impairments (Dias, et al., 1996; Rahman et al., 1997). A study involving adults with psychopathy and the ID/ED Task reports a selective RR impairment (Mitchell et al., 2002). However, a study involving children with psychopathic tendencies found normal RR performance (Blair et al., 2001a).

Procedure: The task is described more fully in Chapter 4. In brief, participants must learn to select between two stimuli presented to them on a computer screen based on feedback provided on the display (the words "correct" or "incorrect"). The stimuli involve up to two dimensions (object shape and line shape). With the exception of the ED shift component, the correct stimulus for each pair is always specified by the shape dimension regardless of the line over which it is superimposed. On each trial, the two test stimuli appear randomly in two of four rectangles positioned towards the perimeter of the screen. Participants used a mouse to make selections. The task consists of nine stages presented in fixed order (to assess the four capacities described above), and participants were required to demonstrate having learnt the given discrimination before proceeding past a given stage. In order to meet the discrimination learning criterion, participants

needed to choose the correct stimulus on eight consecutive trials. The dependent variable was the number of errors made before successfully advancing to the next stage (calculated by the computer). The task was terminated in the event that a subject made greater than 15 errors on a single stage.

Results: Table 6.6 shows the performance of each group on the ID/ED Task. C.L.'s performance on this task was well within or better than the range and CI for the psychopathic, forensic, and community comparison groups. Indeed, he performs at ceiling; all RR components were conducted without error and his performance on the ED component of the task was clearly unimpaired.

Table 6.6: ID/ED Task results

	Object Discrimination	**ED Set Shifting**	**Object Reversal**
CL	0	1	0
Psychopathic			
M	1.40	9.40	7.60
SD	0.89	6.80	8.14
Range	0 to 2	2 to 15	1 to 17
CI	0.29 to 2.5	0.95 to 17.8	0 to 17.7
Forensic controls			
M	1.20	5.00	1.40
SD	0.84	5.67	1.34
Range	0 to 2	1 to 15	0 to 3
CI	0.16 to 2.2	0 to 12.0	0 to 3.1
Controls			
M	0.00	7.00	0.53
SD	0.00	7.45	0.51
Range	-	0 to 15	0 to 1
CI	-	0 to 16.3	0 to 1.2

Task 6.5: The One-Pack Card Playing Task (Newman et al., 1987). A second reversal learning/extinction task was used. Reports in the literature suggest that individuals with developmental psychopathy are impaired on the One-Pack Card Playing Task (Newman et al., 1987).

Procedure: The task procedure is described more fully by Newman and his colleagues (1987). In brief, participants make selections from a single deck of cards. Initially, all selections from the deck are positively reinforced (participants receive ten points). As the task progresses, however, the number of punishments steadily increases by 10% for every 10 cards played. The task continues until the participant elects to stop. Ideally, participants should terminate responding before the probability of punishment exceeds the probability of reward.

Results: In Table 6.7, the performances of C.L. and comparison groups are shown. C.L. terminated the experiment after thirty-eight selections, which is within the range and CI of forensic and community comparison groups, but the number of cards played (indexing sensitivity to changes in reinforcement contingency) was less than the range and CI of psychopathic inmates. Consequently, no evidence for impairment on this task exists.

Table 6.7: One-pack Card Playing Task results

	Cards Selected	Points
CL	38	260
Psychopathic		
M	78.40	232.00
SD	18.53	85.53
Range	58 to 98	120 to 320
CI	55.4 to 101.4	126.1 to 337.9
Forensic controls		
M	56.80	260.00
SD	21.42	43.01
Range	26 to 85	200 to 300
CI	30.2 to 83.4	206.6 to 313.4
Controls		
M	53.40	262.00
SD	19.46	55.86
Range	23 to 73	170 to 310
CI	29.2 to 77.6	192.6 to 331.4

6.5.3: Emotional expression processing

Task 6.6: The Emotional Expression Multimorph Task. Neuroimaging and lesion studies indicate that dissociable neural substrates are crucial for the processing of distinct emotional expressions. Accordingly, the most common expression recognition deficit found in patients with amygdala lesions is for fearful stimuli (Fine & Blair, 2000). Similarly, damage to the VL/OFC regions also results in impoverished emotional expression recognition (Hornak et al., 1996; Blair & Cipolotti, 2000) and imaging work has implicated VLPFC activation to angry facial expressions (Blair et al., 1999). Previous studies utilising the Emotional Expression Multimorph Task report impaired expression recognition in children with psychopathic tendencies (Blair et al., 2001a) and adults with psychopathy (Blair et al., 2004b).

Procedure: This task is a variation of the task designed by Murray and her colleagues (unpublished manuscript) and is used to assess the level of expression intensity required before successful expression recognition is obtained. The stimuli used are taken from the empirically validated Pictures of Facial Affect Series (Ekman & Friesen, 1976). The series consists of stimuli depicting six basic emotional facial expressions (happiness, surprise, fear, sadness, disgust, and anger) that are consistently recognized across a range of different cultures. The individual stimuli were prepared by blending a photographic quality picture of a prototypical expression (i.e., 100% expressed) in varying proportions with the corresponding neutral affect (i.e., 0% expressed). For each continuum, the participant viewed each face as it gradually changed through twenty morphed sequences in 5% increments into one of six prototypical expressions. In brief, participants were instructed to watch as the expression changed, and to report verbally what emotion was being indicated as soon as possible without merely guessing. Participants did not receive any feedback and were free to change their response up until the end of the presentation. After a practice block consisting of one example from each of the six expressions, the participants were presented with 18 test stimuli in random order. Each stage in the morphed sequence was presented for 3 seconds. Participants were scored according to the number of stages (expression intensity) necessary before successful expression recognition occurred (maximum 21). The final answer was scored as the valid response, and in the event of an error, the stage was coded as one point greater than the highest possible score (21 + 1) as a conservative estimate of impairment.

Results: As predicted, C.L. showed a prominent impairment in the recognition of facial expressions of emotion. As reflected by his mean score being outside the range and CI of both forensic and community controls, he suffers from a general facial expression recognition impairment. Inspection of means for individual emotions reveals that he is

outside of the range and CI of forensic controls for fearful, happy and angry facial expressions and beyond the CI for disgusted facial expressions. Furthermore, C.L. scored outside the range and CI of the psychopathic individuals in sensitivity to anger, and beyond the CI for the recognition of happy facial expressions. C.L.'s performance on each of the emotions is beyond the range and CI of that of the community comparison group for every category with the exception of sad and surprised expressions. In short, C.L. shows profound deficits in the recognition of emotional facial expressions.

Table 6.8: Multimorph Expression Recognition Task Number of Stages Required for Recognition

	Fearful	Sad	Happy	Surprised	Disgusted	Angry	Total
CL	17.33	15.00	14.33	15.00	21.33	21.00	17.32
Psychopathic							
M	17.80	15.27	9.40	14.67	16.80	14.27	14.70
SD	3.02	1.94	3.42	4.28	4.69	4.11	2.44
Range	13.33 to 21.67	12.67 to 17.33	5.33 to 14.33	9.00 to 19.67	10.00 to 22.00	7.33 to 17.67	11.39 to 17.00
CI	14.0 to 21.6	12.9 to 17.7	5.2 to 13.6	9.4 to 20.0	11.0 to 22.0	9.2 to 19.4	11.7 to 17.7
Forensic controls							
M	12.93	15.27	8.73	12.07	15.73	11.93	12.52
SD	1.79	1.94	3.01	3.59	4.46	4.21	2.39
Range	10.67 to 15.67	12.67 to 17.33	5.00 to 13.33	9.00 to 19.67	10.00 to 22.00	7.33 to 17.33	9.28 to 15.72
CI	10.7 to 15.2	8.5 to 19.0	5.0 to 12.5	7.6 to 16.5	10.2 to 21.3	6.7 to 17.2	9.6 to 15.5
Controls							
M	13.80	13.93	6.73	16.07	14.20	13.93	13.11
SD	2.30	2.92	2.42	4.59	4.68	4.02	2.62
Range	11.33 to 16.67	9.33 to 17.33	4.33 to 10.33	12.00 to 22.00	8.67 to 21.67	8.67 to 17.33	9.17 to 16.28
CI	10.9 to 16.7	10.3 to 17.6	3.7 to 9.7	10.37 to 21.76	8.39 to 20.0	9.0 to 16.4	9.9 to 16.4

Task 6.7: Audio Emotion Recognition (Scott et al., 1997). The Vocal Affect Recognition Test assesses recognition of emotion conveyed through prosody. On each of the 65 test trials, one of six bi-syllabic concrete nouns (digitalized at 22kHz) of neutral denotation (carpet, finger, hammock, motor, sailor, daughter) are presented in random order from a Macintosh G3 laptop computer.

The test consists of 65 neutral words (e.g., carpet) spoken by native English speakers (3 male and 3 female) that convey emotions of happiness, disgust, anger, sadness and fear (13 examples of each emotion). Participants were asked to listen to each stimulus and to identify the emotion that the speaker was most likely to have been feeling based on how the word was spoken. The result is a score out of 12 for each basic vocal emotion providing an index of sensitivity to emotion presented in the auditory modality.

Results: In striking contrast to C.L.'s recognition of facial expressions of emotion, he performed within the range and CI of both the forensic and community comparison groups for the recognition of auditory emotional expressions. Furthermore, he showed superior performance to individuals with developmental psychopathy for the identification of distress cues (fearful and sad vocal affect).

Table 6.9: Auditory Emotional Expression Recognition Task results (Errors)

	Fearful	Sad	Happy	Disgusted	Angry	Total
CL	0	1	7	2	3	2.60
Psychopathic						
M	5.00	4.20	4.20	2.80	3.80	4.00
SD	2.00	1.48	2.58	2.95	2.17	1.57
Range	3.00 to 8.00	2.00 to 6.00	1.00 to 8.00	0.00 to 7.00	2.00 to 7.00	2.20 to 5.40
CI	2.5 to 7.5	2.4 to 6.0	1.0 to 7.4	0.0 to 6.5	1.1 to 6.5	2.1 to 5.9
Forensic controls						
M	1.80	2.60	4.60	2.00	2.60	2.72
SD	1.10	1.15	3.36	0.71	1.52	0.30
Range	0.00 to 3.00	1.00 to 4.00	1.00 to 10.00	1.00 to 3.00	1.00 to 4.00	2.40 to 3.20
CI	0.44 to 3.2	1.2 to 4.0	0.4 to 8.8	1.1 to 2.9	0.7 to 4.5	2.3 to 3.1
Controls						
M	1.60	1.20	5.00	3.60	2.20	2.72
SD	1.67	0.84	2.55	2.19	1.79	1.24
Range	0.00 to 4.00	0.00 to 2.00	3.00 to 9.00	1.00 to 7.00	0.00 to 4.00	1.40 to 4.40
CI	0 to 3.7	0.2 to 2.2	1.8 to 8.2	0.9 to 6.3	0.0 to 4.4	1.2 to 4.3

6.5.4: Social cognition

The following tasks were aimed at assessing two main aspects of social cognition. First, C.L.'s ability to infer internal mental states in others (tasks 6.8 and 6.9) was examined. Second, the ability to assess the appropriateness of a set of behaviours presented in a variety of social contexts was investigated (task 6.10). The ability to infer mental states such as thoughts, feelings, or intentions in others and to predict and understand behaviour on the basis these mental states has been referred to as "theory of mind" (Premack & Woodruff, 1978). Impairments of theory of mind have been implicated in autism spectrum disorders (Happé, 1994; Baron-Cohen et al., 2000) and schizophrenia (Corcoran, Mercer, & Frith, 1995; Frith & Corcoran, 1996). The neural structures implicated in theory of mind are the amygdala, superior temporal sulcus, anterior paracingulate, and the OFC (Baron-Cohen et al., 1994; Baron-Cohen et al., 2000; Gallagher et al., 2000; Fine, Lumsden, & Blair, 2001; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). Studies investigating theory of mind in individuals with psychopathy find no evidence of impairment (Blair et al., 1996; Blair & Cipolotti, 2000; Richell et al., 2003).

Task 6.8: The Joke Comprehension Test (Corcoran et al., 1997). C.L. completed the Joke Comprehension Test (Corcoran et al., 1997), which contains 10 jokes that are understood only with reference to mental states of the characters and 10 jokes that can be understood through physical and semantic analysis. C.L. and healthy age and IQ matched controls from the community were asked to explain each cartoon and to rate how humorous they found them on an ascending scale of 1 to 10. Their explanations were recorded verbatim and scored following the testing session. A score of 1 was given for

each mental state cartoon that was appropriately explained with reference to mental states. Similarly, each response that made reference to the physical situation was given a score of 1. Individuals with deficient theory of mind show poor comprehension of the mental state cartoons, but normal comprehension and humour to the physical cartoons (Corcoran et al., 1997).

Figure 6.2: Cartoons from the Joke Comprehension Test (Corcoran et al., 1997)
A) Theory of mind cartoon. Theory of mind facilitates the comprehension of this cartoon: to understand the joke, participant must realize the dog-owner's erroneous assumption about the size of the rabbit.

"Go on boy, rabbits, get 'em!"

B) Physical state cartoon. This cartoon can be understood without reference to mental states.

"It's really no wonder that we're endangered"

Results: C.L., in making only one error, was able to explain the nature of each cartoon with reference to both physical and mental states for each sample. His humour rating for each cartoon is equivalent for both physical and theory of mind samples signifying a comparable degree of comprehension and appreciation. His scores are well within the range of the comparison group; C.L. shows no evidence of theory of mind impairments on the task.

Table 6.10: Performance on the Joke Comprehension Test

	Theory of mind comprehension (max = 10)	**Theory of mind comic rating (max = 10)**	**Physical comprehension (max = 10)**	**Physical comic rating (max = 10)**
CL	9.00	4.50	10.00	5.60
Controls				
M	9.00	4.76	9.20	4.58
SD	1.00	1.86	1.79	1.93
Range	8 to 10	2.30 to 6.70	6.00 to 10.00	1.90 to 6.40
CI	7.8 to 10.2	2.4 to 7.1	7.0 to 11.4	2.2 to 7.0

Task 6.9: The Reading the Mind in the Eyes Task (Advanced Theory of Mind; Baron-Cohen, Wheelwright, Hill, Rate, & Plumb, 2001). The Eyes Task investigates an individual's ability to infer mental states of others based on visual information from photographs depicting a person's eyes. Previous reports have indicated that high-functioning individuals with autism-spectrum disorder and above average IQ show impaired performance relative to matched control participants (Baron-Cohen et al., 1997). An imaging study has implicated the amygdala in processing the images.

Procedure: Details about the development of this task are presented in a previous study (Baron-Cohen et al., 1997). In brief, participants are presented with 36 photographs depicting a small portion of an actor's face including the eyes, eyebrows and the eyes of

actors and the region of the face around the eyes. Four complex mental state descriptors (e.g. dispirited, bored) were printed around the photo, one at each corner. One of these words (the target) correctly identified the mental state of the person in the photo, whilst the others were included as foils (e.g., annoyed, hostile, horrified, or preoccupied). Participants are asked to choose the word that best describes what the person in the picture is thinking or feeling. Definitions for each of the words were made available to each participant.

Results: In contrast to C.L.'s performance on basic emotional expression recognition, his performance on the task is within the upper range of normal. Therefore, the data do not support suggestions of an association between acquired sociopathy and theory of mind deficits.

Table 6.11: Reading the Mind in the Eyes Task results

	Correct Responses (max = 36)
CL	30
Psychopathic	
M	25.20
SD	4.71
Range	19.00 to 31.00
CI	19.3 to 31.1
Forensic controls	
M	25.00
SD	4.30
Range	20.00 to 31.00
CI	19.7 to 30.3
Controls	
M	29.40
SD	4.16
Range	23.00 to 33.00
CI	24.2 to 34.6

Task 6.10: The Social Situations Task (Dewey, 1991). The Social Situations Task indexes an individual's ability to detect situations that are likely to provoke disapproval in others. These include acts for which no formal societal prohibitions exist, but which may provoke anger or aggression in observers. The study includes nine short stories depicting behaviours that are either normative (in line with social norms) or violations (violating social norms). At various points in the story, participants rate the appropriateness of the behaviours described by giving a score from 0 to 3 expressing the severity of the violation (a score of "0" denotes normal behaviour and 3 indicates shocking behaviour). By examining the frequency with which social norms and their violations are identified, the task provides an estimate of an individual's ability to detect socially inappropriate behaviour. Previous research suggests that the task distinguishes between two types of pathologies. Although this task is sensitive to cognitive deficits associated with acquired sociopathy, individuals with developmental psychopathy do not show impaired performance (Blair & Cipolotti, 2000). Individuals with autism also exhibit difficulties with this task; however, they respond abnormally to both normal and inappropriate behaviours (Dewey, 1991). In contrast, the SRR conceptualisation of acquired sociopathy predicts that patients with OFC lesions will rate inappropriate behaviours as appropriate. A previously documented case of acquired sociopathy, patient J.S., showed precisely this pattern of impairment (Blair & Cipolotti, 2000).

Procedure: Scenes describing social situations were presented to the participants on a computer screen and read aloud. At specified points in the stories, participants were asked to verbally rate the behaviour just described to them in the story according to how they thought most people would judge that behaviour had they witnessed it. The following scale was used: 0 = Fairly normal behaviour in that situation; 1 = Rather

strange behaviour in that situation; 2 = Very eccentric behaviour in that situation; 3 = Shocking behaviour in that situation.

Eight social scenes were described to the participant in total. Between 2 and 4 behaviours were rated in each scene. These behaviours were divided into normative scenes (e.g., "when she woke up there was just enough time to dress and get to the airport, so she skipped her breakfast") and violations (e.g., "without wasting any time, even before the introductions, he asked the hostess when dinner would be served"). Mean ratings for these two types of behaviour were calculated. A sample story is presented below:

> *Keith, age twenty-five, was a file clerk who worked in an office in the city. At noon, he took his lunch to a small park and sat on a sunny bench to eat. Often he tore part of a sandwich into bits, scattering it on the ground for pigeons ( ). One day when he came to his favourite bench a baby carriage was parked beside it. Keith noticed that a young woman was swinging an older child nearby. The baby in the carriage began to cry but the mother did not hear this because the swing was squeaking. Now, Keith had learnt that when his baby nephew screamed, sometimes this meant that a pin in his diaper had opened. Rather than bother the mother in the park, Keith quickly checked the baby's clothing to see whether he could feel an open pin ( ).*

Results: C.L. shows significant performance decrements on this task relative to forensic, psychopathic and community groups scoring less than the range and CI of each group. He correctly identified only 5 of 11 social norm violations. This contrasts with his above-average performance on the theory of mind tasks. The results lend support to the idea that although the identification of socially inappropriate behaviours is impaired in individuals with autism (Dewey, 1991; Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999), it can also be impaired in an individual who exhibits no theory of mind deficits. Therefore, although theory of mind deficits may be sufficient for producing deficits in the recognition of socially inappropriate behaviour, it is not the only factor.

Table 6.12: Social Situations Task results

	Social violations Correctly identified (0 to 11)	Normal behaviours labelled violations (0 to 11)	Inappropriateness rating for violations (0 to 33)
CL	5	1	8
Psychopathic			
M	8.40	2.00	13.20
SD	1.14	1.87	3.70
Range	7 to 10	0 to 5	10 to 19
CI	7.0 to 9.8	0 to 4.3	8.6 to 17.8
Forensic controls			
M	8.40	2.20	14.80
SD	0.55	1.64	0.84
Range	8 to 9	0 to 4	14 to 16
CI	7.7 to 9.1	0.2 to 4.2	13.8 to 15.8
Controls			
M	9.80	2.40	18.80
SD	1.30	2.30	5.45
Range	8 to 11	0 to 6	12 to 27
CI	8.2 to 11.4	0 to 5.3	12.0 to 25.6

6.6: Discussion

Following bilateral trauma to the frontal region involving OFC, C.L. presented with a severe disturbance in affective, behavioural, and interpersonal functioning. His aberrant behaviour and interpersonal style bore a strong resemblance to the construct of psychopathy (Hare, 2003); based on numerous detailed clinical accounts, C.L. can be described as a case of "acquired sociopathy" (Damasio, 1994). This chapter presents measures of decision- making, instrumental learning and relearning, emotional expression recognition, and social cognition. C.L.'s performance was compared with a forensic sample of psychopathic and non-psychopathic adults and with a sample of healthy community controls of comparable age and IQ. C.L. showed clear deficits in decision-making, conditional learning, emotional expression recognition, and in the identification

of inappropriate social behaviours. Contrary to expectations, C.L. did not show pronounced reversal learning deficits. Despite showing normal ability to form stimulus-reinforcement associations, C.L. demonstrated clear conditional learning impairment. Finally, no evidence of difficulties with theory of mind was indicated in C.L. or the forensic samples.

6.6.1: Decision-making performance

The data implicating orbital regions of the frontal cortex in decision-making is considerable (Bechara et al., 1994; Bechara et al., 2000). In line with predictions, C.L., along with a group with developmental psychopathy, showed impaired performance on the Iowa Gambling Task relative to both the forensic and community controls. It is important to note that individuals with amygdala dysfunction also show deficits on this task (Bechara et al., 1999). The source of their impairment, however, may differ. Using psychophysiological measures, Bechara and his colleagues (1999) draw a distinction between the impairment shown by patients with amygdala dysfunction from that of patients with frontal lesions. They suggest that patients with lesions of the amygdala fail to avoid disadvantageous selections because of a general insensitivity to reinforcement (they fail to generate skin conductance responses following reward or punishment). In contrast, patients with lesions to the OFC were able to generate skin conductance to reward and punishment, but were unable to generate anticipatory responses when making high-risk deliberations. Although the current chapter shows that both C.L. and individuals with developmental psychopathy were impaired on the Iowa Gambling Task, the underlying neurocognitive deficit contributing to their respective performances may be different.

Regions of the frontal cortex, particularly ventrolateral regions have been implicated in RR (Cools et al., 2002; O'Doherty et al., 2003). One of the dominant theoretical frameworks of VL/OFC function posits that the system is crucial for the representation of, and detection of changes in, the reinforcement value of stimuli (Rolls, 1996; Rolls, 2004). In the current set of experiments, the performance of C.L. and a forensic sample was investigated on two tasks thought to be sensitive to reversal learning and extinction: the ID/ED Task and the One-Pack Card Playing Task. Contrary to expectations, C.L. showed no evidence of reversal learning or extinction impairment. Individuals with psychopathy show deficits on both tasks (Newman et al., 1987; Mitchell et al., 2002). The ID/ED Task features an abrupt change in reinforcement contingency that may be less sensitive to more subtle reversal learning impairments. The One-Pack Card Playing Task, however, contains a subtle shift in reward contingency, and C.L.'s performance was within the range and CI for the comparison groups and was better than that of individuals with psychopathy. This might suggest that despite the extent of the lesion, other more general factors, such as IQ, may act as compensatory mechanisms to facilitate performance on this task. Taken together, the data presented in this chapter do not provide any evidence of a reversal learning impairment for C.L.

6.6.2: Dissociable forms of instrumental learning

This study examined two forms of instrumental learning. The Passive Avoidance Task provides a measure of stimulus-reinforcement learning and the Ask for Money Task features conditional learning. Crucially, the two forms of learning are dissociable at the neural level. The amygdala is necessary for stimulus-reinforcement but not conditional learning (Baxter & Murray, 2002). In contrast, dorsolateral regions play a crucial role in conditional, but not stimulus reinforcement learning (Petrides, 1985a; 1985b; 1990;

1997). For example, monkeys with lesions to the periarcuate region of dorsolateral prefrontal cortex are unable to complete a conditional discrimination; they are unable to produce one response when stimulus A is present and a second response when stimulus B is present for reward (Petrides, 1985a). The same monkeys, however, were able to respond to one stimulus for reward and to withhold responding to another stimulus that was associated with non-reward. Given the extent of C.L.'s frontal lesions, and intact temporal lobes, it was predicted that he would show deficient conditional learning, but intact passive avoidance learning. The results supported this prediction. The experiments show a double dissociation between acquired and developmental psychopathy. Although C.L. showed intact passive avoidance learning and diminished conditional learning, the opposite pattern of impairment was found in the group with developmental psychopathy. The result provides additional support for the notion that dissociable forms of instrumental learning exist.

6.6.3: Emotional expression recognition

Emotional expression processing has been linked to the OFC (Hornak et al., 1996; Blair & Cipolotti, 2000) and a more specialized role for the amygdala particularly with fearful (Breiter et al., 1996; Morris et al., 1996; Phillips et al. 1997; Phillips et al., 1998), sad (Blair et al., 1999), and happy (Breiter et al., 1996) facial expressions. In addition to being impaired relative to the healthy community sample, C.L. performs outside the range of forensic controls and even psychopathic individuals. C.L. shows evidence of a general facial expression recognition impairment, which was perhaps most pronounced for angry and disgusted facial expressions. This result replicates a previous finding (Hornak et al., 1996), and based on the severity of the angry expression recognition deficit, provides a partial replication of Blair and Cipolotti's (2000) study involving J.S., a similar case of

acquired sociopathy. Relative to a forensic comparison group, J.S. showed deficient recognition of happiness, anger, disgust, and sadness.

Interestingly, C.L.'s performance did not cross modalities; he performed close to ceiling on the auditory emotional expression identification task. It should be noted that at least two studies have found vocal expression recognition deficits in patients with bilateral OFC damage (Hornak et al., 1996; Hornak et al., 2003). However, C.L. may not be unusual in this regard. It has been noted that emotional expression recognition deficits do not necessarily cross modalities even among frontal patients (Hornak et al., 2004; Rolls, 2004).

6.6.4: Theory of mind

It has been suggested that a circuit involving the OFC and amygdala is involved in making inferences about intentions, thoughts, and feelings of others—a process collectively referred to as "theory of mind" (Brothers & Ring, 1992; Stone et al., 1998; Baron-Cohen et al., 2000; Sabbagh, 2004). At least three studies have investigated theory of mind performance in psychopathic individuals diagnosed by the PCL-R; however, none have found evidence of impairment (Blair et al., 1996; Blair & Cipolotti, 2000; Richell et al., 2003). Despite the sensitivity of the tests used, the present investigation did not indicate theory of mind dysfunction in the forensic samples or in C.L. The study involving J.S. also failed to find evidence for difficulties representing mental states in others (Blair & Cipolotti, 2000). Furthermore, suggestions of links between theory of mind deficits and aggression have been sharply criticized (Blair, 2003).

Despite these findings, at least one study has reported theory of mind deficits in patients with lesions to the OFC. Stone and her colleagues (1998) report that patients with lesions involving the orbital, but not dorsolateral regions of the frontal cortex are

impaired on the Social Faux Pas Test. This task consists of a series of stories depicting social exchanges in which the speaker makes an inappropriate remark. The authors interpret a failure to identify the social faux pas as evidence of theory of mind dysfunction. However, the task is arguably less a measure of theory of mind than it is a measure of the identification of inappropriate social behaviour. Indeed, in the same study, the patients with lesions to OFC did not differ from comparison individuals in their judgement of first and second order theory of mind problems. Furthermore, patients with OFC lesions were able to identify how the victims of the social faux pas would have felt and why the speaker said what they said. Responding correctly to these questions is arguably a more sensitive index of mental state representation than the identification of inappropriate social behaviour.

6.6.5: Identification of social violations

The present investigation used Dewey's Social Situations Task (1991) to explore C.L.'s ability to detect inappropriate social behaviour. Neuroimaging and patient studies suggest that the OFC plays a role in monitoring and adjusting social behaviours (Blair & Cipolotti, 2000; Berthoz, Armony, Blair, & Dolan, 2002). The ability to make such judgements is thought to be an index of the SRR system (Blair & Cipolotti, 2000; Blair, 2004). The SRR system is activated by negative valenced expressions (particularly anger and disgust) and situations associated with social disapproval. The SRR system is thought to be functionally dissociable from the system that computes or codes changes in reinforcement (Blair & Cipolotti, 2000; Blair, 2004). The results of the present investigation provide support for this conceptualisation: C.L. showed a profound impairment in the recognition of socially inappropriate behaviour. In addition, he showed a striking deficit for identifying emotional facial expressions, particularly anger and

disgust. In contrast, C.L. did not exhibit an equivalent level of impairment on the reversal component of the ID/ED Task, or the more subtle shift in reinforcement value on the One-Pack Card Playing Task. Both of these tasks are sensitive measures of reversal learning deficits in adults with developmental psychopathy (Mitchell et al., 2002).

6.7: Conclusion

The current investigation compared the performance of a patient with acquired sociopathy to individuals with developmental psychopathy as well as to forensic and community controls. In line with predictions, C.L. showed a strikingly different pattern of impairment than developmental psychopaths. Specifically, he showed a collection of deficits that are associated with an impaired SRR system, only subtle, if any, impairment of the RR system, and a decision-making impairment. Conversely, individuals with developmental psychopathy showed no evidence of SRR dysfunction, but did exhibit impaired RR and decision-making. Contrary to OFC accounts of theory of mind, C.L. showed no evidence of impairment on two tasks that are considered to be highly sensitive to making inferences about mental states. With respect to instrumental learning, a double dissociation was found. C.L. showed pervasive conditional learning deficits in the absence of passive avoidance learning deficits. Conversely, developmental psychopaths showed intact conditional learning and pronounced passive avoidance deficits. Together, the results provide additional support for the notion that SRR deficits can be associated with reactive aggression in the absence of evidence for theory of mind or pronounced RR dysfunction. It also provides further evidence that acquired sociopathy and developmental psychopathy are dissociable at a neurocognitive level. The different forms of impairment may underlie the two different forms of aggression, reactive versus

instrumental, that characterise each acquired sociopathy and developmental psychopathy respectively.

Chapter 7

fMRI Investigation of Response Reversal Versus Value Change

7.1: Introduction

The neural basis of instrumental relearning and stimulus devaluation has been the focus of a considerable body of research. Regions of the VL/OFC cortex are implicated in altering responding to changes in reinforcement (Rolls, 1999). Theories emphasize the role of the OFC in response change and in flexibly encoding the value of stimuli (Rolls, 1999; Rolls, 2004). These theories often implicitly subsume the process of stimulus value encoding, inhibition of a prepotent response, and response change within the VL/OFC. However, the precise neural loci of these potentially dissociable functions remains unclear. This chapter presents an event-related fMRI study investigating neural activation to changes in the value of reinforcement together with, and in the absence of, demands for response change. The experiment was designed to enable contrasts that would distinguish between processes involved in detecting changes in the value of reinforcement from processes involved specifically with response change. The aim was to address the continuing question of whether the neural response to changes in valence or magnitude of reinforcement is dissociable from activation associated with response change.

Studies emphasizing flexible encoding of reward value began with experiments involving the response of non-human subjects to primary reinforcement. These investigations suggest that the OFC is involved in encoding the reinforcement value of both pleasant and aversive stimuli. For example, single unit recording of neurons in the caudolateral region of the macaque OFC identified neurons that respond to the taste of a stimulus with nutritive value, such as fruit juice, but not to non-nutritive tastes such as

quinine (Rolls et al., 1990). Neurons in the OFC selectively fire to an olfactory stimulus associated with food, and decrease firing when that food is devalued (Critchley & Rolls, 1996). Single unit recordings of neurons within the OFC of behaving rhesus monkeys reveal associative stimulus-specific activity (Thorpe, Rolls, & Maddison, 1983). Thus, in a visual discrimination task, a percentage of neurons responded selectively to a stimulus associated with reward, and other neurons to the stimulus associated with punishment. Interestingly, when the reinforcement associations were reversed, the pattern of neuronal firing reversed accordingly.

Later studies used imaging techniques to investigate neural responding to reinforcement in human subjects. These studies provide further support for the role of the OFC in encoding the value of primary reinforcement. For example, activation has been reported in adjacent regions of VLPFC to sucrose and salt (O'Doherty et al., 2001b). A second study reported activation of the ventral striatum and OFC to the expectation of receiving sucrose (O'Doherty, Deichmann, Critchley, & Dolan, 2002). Furthermore, in a human analogue to the study involving macaque monkeys, a recent investigation showed that OFC responding to a food odour decreased when the food associated with the odour was eaten to satiety (O'Doherty et al., 2000).

The OFC is also involved in encoding abstract forms of reinforcement. For example, participants engaging in a simple motor response task showed activity in the inferior frontal left amygdala, inferior frontal gyrus, left hippocampus, and right VLPFC to positive versus negative verbal feedback (Zalla et al., 2000). Thut and his colleagues (1997) report greater activation for monetary versus verbal reinforcement in dorsolateral (BA 10 and 44) and VLPFC (BA 47) in addition to thalamic and midbrain activity. In a probabilistic object discrimination and reversal task involving monetary reward, O'Doherty and his colleagues report medial frontal activity to reward and lateral activity

to punishment (O'Doherty et al., 2001a). The activity in the medial and lateral prefrontal cortex correlated with the magnitude of reward and punishment respectively.

The evidence implicating VL/OFC regions of the frontal cortex in processing both primary and abstract reinforcement is compelling. Data also exist suggesting a special role for at least some regions of the prefrontal cortex with respect to response suppression and reversing a previously rewarded response. For example, patients with lesions to the OFC show difficulties reversing a previously rewarding response (Rolls et al., 1994; Fellows & Farah, 2003; Hornak et al., 2003; Berlin et al., 2004; Hornak et al., 2004). Bechara and his colleagues (2000) have shown that individuals with damage to ventromedial regions of frontal cortex are unable to adjust their immediate behaviour for future reward. In a study conducted by Cools and her colleagues (2002), VLPFC was activated specifically by probabilistic errors that preceded response alteration. Imaging studies consistently implicate this region in RR (Rogers et al., 2000; Clark et al., 2004; Cools et al., 2004). The data therefore suggest a specific role for the region in response change. Evidence, however, also suggests that VL/OFC is involved in the inhibition of a prepotent response. For example, using a detour reaching task, Wallis and his colleagues showed that monkeys with lesions involving the OFC encountered difficulties inhibiting the dominant, but unsuccessful, direct reach response (Wallis, Dias, Robbins, & Roberts, 2001). In humans, VLPFC activation has been reported for no-go trials of a go/no-go task (Liddle, 2001).

Based on these findings involving response change and response inhibition, researchers have speculated that discrete regions within the VL/OFC may be performing separate functions during response change. For example, O'Doherty and his colleagues (2001a) have suggested that a valence-dependent functional dissociation exists between the orbital and ventral regions of the prefrontal cortex. Based on their fMRI data, they

propose that medial regions are specifically involved in processing reward and lateral regions in processing punishment. Similarly, it has been suggested that medial regions of OFC have also been implicated in processing the reward value of response options when correct choice is unclear and requires the integration of information over several trials to determine (Elliott, Rees, & Dolan, 1999; Elliott, Dolan, & Frith, 2000). In contrast, the lateral regions of the frontal cortex are thought to be recruited in circumstances for which the inhibition of a previously rewarded (or otherwise prepotent) response is necessary (Elliott et al., 2000). Indeed, an fMRI investigation of a delayed matching and non-matching to sample task showing medial activation to matching and lateral activation to non-matching provides support for this position (Elliott & Dolan, 1999). Thus, the medial regions (BA 24 and 25) were involved when participants were instructed to select the familiar pattern in a pair. In contrast, ventrolateral regions (BA 11) of frontal cortex were involved when participants had to inhibit selecting the familiar stimulus in favour of the unfamiliar one. This formulation offers a reasonable account of why lesions to either medial (inconsistent selection of rewarding stimuli) or lateral (perseverating on unrewarded stimuli) regions of the prefrontal cortex can lead to impaired decision-making. Interestingly, in a study involving patients with OFC lesions, Berlin and colleagues report both a reduced tendency to shift responding following a large punishment *and* a failure to maintain responding following a large reward in patients with lesions involving the OFC (Berlin et al., 2004). Crucially, eight of the ten patients in the study had lesions involving both medial and lateral prefrontal regions.

These studies suggest a potential dissociation of function within the frontal lobes based on valence of reinforcement or demands for response change. However, the studies designed to investigate response change and changes in reinforcement value are often confounded; the tasks that are commonly used may not adequately distinguish modulated

neural activity to changes in reinforcement value from neural responding due to increased demands for response change. For example, decision-making tasks typically involve reversing a response from a previously rewarding stimulus to a previously punished stimulus. Behavioural change of this kind, however, may require several stages of processing. For example, RR may entail: (1) detecting a contingency change; (2) evaluating whether the magnitude of the contingency change is sufficient to warrant behavioural change; (3) inhibiting the previously relevant response; and (4) initiating a newly rewarded response. Thus, RR may require several stages of processing mediated by different regions within the prefrontal cortex. To date, almost all investigations treat reversal learning as a unitary construct.

However, at least one other study has attempted to disentangle some of the potentially dissociable components of RR (O'Doherty et al., 2003). A two-condition probabilistic reversal learning paradigm was used. In the "choice" condition, participants chose between two fractal images (one associated with higher monetary rates of reward, and the other with punishment). In a second, "imperative" condition, the computer generated the selections while participants were asked simply to "track" the currently rewarded stimulus. It was thought that comparisons between the choice and imperative conditions would distinguish neural responding associated with response change from neural activity associated with encoding changes in the valence of reinforcement. The authors implicate medial regions in representing outcome, particularly reward. They found that increased activation in medial and left lateral OFC was associated with feedback that preceded maintaining the same response on the next trial as opposed to feedback that preceded response change. Finally, parts of medial and central OFC were significantly more active during the choice condition than the imperative condition, thus suggesting a role that also includes decision-making. Activity in the anterior insula and

caudolateral OFC was associated with punishment that preceded a change in response relative to punishment that did not; however, this was found in both the choice and imperative conditions, which contradicts interpretations that emphasize involvement of this region strictly in terms of response change. The authors propose that RR deficits seen following injuries to this region are not caused by a failure to inhibit a prepotent response, but rather, by a failure to detect changes in reinforcement value. One region that appeared to be selectively involved in response change in the choice over imperative condition was the dorsal region of anterior cingulate. They suggest a link between this region and autonomic facilitation of response change.

The complexity and apparent overlapping nature of the results may reflect confounds that still exist in the experimental design. That is, the experiment may not have been entirely successful in its attempts to dissociate neural responding to changes in valence from that related to response change. For example, in the Probabilistic Response Reversal Task, two distinct forms of feedback are indistinguishable. Probabilistic errors (error feedback received after choosing the correct stimulus) are indistinguishable from error feedback that represents a true reversal in reinforcement associations. In the case of a probabilistic error, overriding error feedback to maintain the same response may activate similar regions as actual response change. Alternatively, neural firing in a region of frontal cortex may be required to reach a certain threshold before response change is initiated. In either case, a considerable degree of overlap may exist in neural activity whether or not an error precedes response change.

The following investigation attempts to distinguish further neural responding to changes in reinforcement value from that related to initiating response change. The aim was to determine whether OFC activation classically viewed following RR is specific to response change, or generalises to all changes in stimulus value, regardless of demands

on responding. A novel conditional learning task is presented in which participants make one of two responses to multiple stimuli on a screen. Following an initial acquisition phase, the value of the reinforcement changed for four of the five stimuli. The nature of the change for each stimulus varied according to a set schedule: (1) RR with valence change; (2) RR with only magnitude change (the valence remained constant); (3) changes in the valence of the reinforcement without demands for response change; (4) changes in the magnitude of reinforcement without changes in valence and without demands for response; and (5) a control condition in which valence and magnitude remain unchanged. The results are discussed within the context of current models of VL/OFC function.

7.2: Methods

7.2.1: Participants

Thirteen right-handed healthy volunteers (4 females and 9 males) took part in the study, which was approved by the National Institute of Mental Health Institutional Review Board. All participants underwent a physical exam by a physician, were free of psychotropic medication, and were screened to exclude those with a history of psychiatric or neurological illness.

7.2.2: fMRI data acquisition

Subjects were scanned during task performance using a 1.5 Tesla GE Signa scanner. Two-hundred seven functional images were taken with a gradient echo-planar imaging (EPI) sequence (repetition time = 2500ms, echo time = 40ms, 64 x 64 matrix, flip angle 90°, FOV 24cm). Coverage was obtained with 29 axial slices (thickness, 4-mm; in-plane resolution, 3.75 x 3.75 mm). A high resolution anatomical scan (three-dimensional Spoiled GRASS; repetition time = 8.1ms, echo time = 3.2ms; field of view =

24cm; flip angle = 20°; 124 axial slices; thickness = 1.0 mm; 256 x 256 matrix) in register with the EPI dataset was obtained covering the whole brain.

7.2.3: The Stock Market Task

The Stock Market Task is a novel task designed to engage the participant in a decision-making task with changing reinforcement values and response demands. Participants completed a novel stock exchange task in which they were asked to decide which stock to sell or to retain based on the market values presented on a computer screen. The reinforcement values of learned responses were unexpectedly altered; the alteration resulted in varied demands for response change with and without changes in valence of reinforcement, and changes in reinforcement without demands for response change. The participants received the following instructions:

> *You are a successful trader on Wall Street who controls stock for very wealthy clients. Unfortunately, the stock market has recently crashed and things are a mess. Some stocks have lost value and will continue to lose value. Other stocks have lost value, but will bounce back and are best kept. Still others were unaffected by the crash. Of these, some might drop if kept, but others will continue to grow. Your clients are turning to you to save their fortunes. It's up to you to decide which stock are bad and should be sold immediately and which stock are good and are best to keep around for awhile.*
>
> *Each individual stock item will appear as a black and white drawing depicting the goods that the stock represents. Next to the stock will be a square indicating how much money you will lose or gain relative to its original pre-crash value if you sell right away. If you choose to sell, this dollar figure is how much you will gain or lose. Otherwise, you may choose to wait. Once you make your selection, you will be told how much money your selection won or lost you, as well as how much money you would have made had you made the opposite choice. The dollar figure is always expressed relative to the pre-crash value of the stock.*

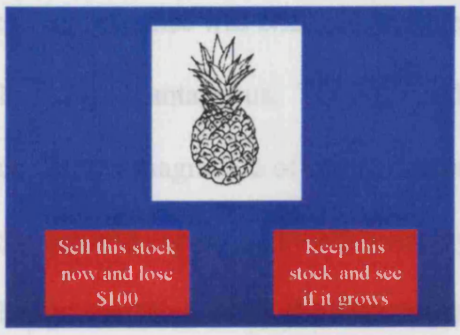
Stocks will appear one at a time. If you wish to sell the stock shown, please press "1." If you wish to keep the stock for a while, please press "2." Once you have made your decision, the results will appear in two red boxes: One showing you the result of your choice and the other showing you what the result would have been had you chosen the other option (selling instead of keeping or vice versa). It is VERY important that you pay attention to both boxes to make sure that you are making the right choice each time.

Remember, to keep your clients happy you need to learn which stock items are bad and are best sold and which are good, and are best kept.

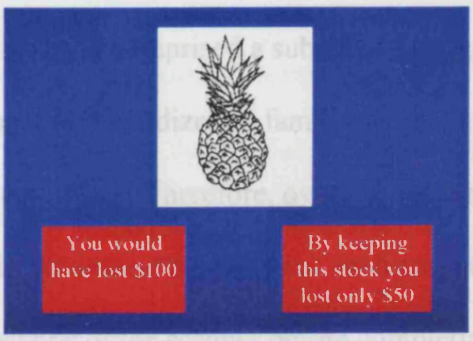
On each trial, participants had the option either to sell the stock immediately, or to retain it to determine whether it grows. Initially, participants learned that for each stimulus, the best outcome was associated with retaining the stock rather than selling it immediately. This "acquisition phase" consisted of 15 trials in each run (3 presentations of each stimulus). The entire experiment consisted of 8 runs and 120 acquisition stimuli. Each stimulus was associated with a unique level of reinforcement for retaining the stock. Figure 7.1(A), 7.1(B), and 7.1(C) illustrate sample response option screen, reinforcement screen, and trial schematic.

Figure 7.1: Sample user interfaces

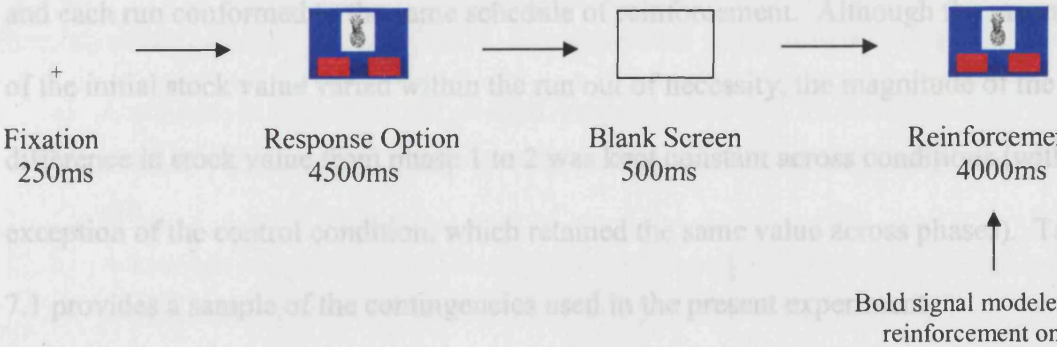
A) Response option screen: The stock product was depicted by a Snodgrass image (a pineapple below). The current market value of the stock is shown in a red box on the left. In this case, the value of the stock is set at $100 less than the pre-market crash value. Thus, selling the stock immediately would result in a net loss of $100. Participants also had the option of retaining the stock, which is indicated in the red box on the right of the figure, which displays the option, but not the value of that option.



B) Reinforcement screen: The reinforcement screen displayed the consequence of the response, but also what the consequence would have been had the participant made the opposite choice. This ensured that participants were exposed to the reinforcement value of each response on every trial.



C) Trial schematic



Following the acquisition phase, the reversal phase occurred without warning or perceivable change. The reversal phase consisted of five presentations of each stimulus for each run (25 trials in total). During the reversal phase, two out of the five response contingencies were reversed. For the RR and valence change conditions (RRVC), the valence of the previously correct response was changed and the previously disadvantageous response became advantageous. The RR condition stimuli (RR), in contrast, had a fixed valence, but the magnitude of reinforcement changed such that the previously incorrect response became correct and the previously correct response became incorrect. In the valence change condition (VC), the valence of reinforcement was reversed, but the advantageous response remained the same. In the magnitude change condition (MAG), the stimulus changed in value, but retained the same valence and response. In the control condition, the value and response options remained unchanged (CON). The stimuli that were used comprised a subset of Snodgrass and Vanderwart's (1980) black and white images (standardized in familiarity and visual complexity) randomly assigned to each condition. Therefore, over the course of eight runs, participants saw 200 trials in the second phase (40 of each condition). Each participant underwent a practice task outside of the scanner before completing eight different versions of the task while being scanned. A different set of stimuli were used in each run, and each run conformed to the same schedule of reinforcement. Although the magnitude of the initial stock value varied within the run out of necessity, the magnitude of the difference in stock value from phase 1 to 2 was kept constant across conditions (with the exception of the control condition, which retained the same value across phases). Table 7.1 provides a sample of the contingencies used in the present experiment.

Table 7.1: Sample reinforcement table

Condition	Sell Now Value	Reinforcement for "Keep Stock" Response	
		Phase 1	Phase 2
RRVC	- $75	$50	- $100
RR	- $100	- $50	- $200
VC	- $100	- $50	$100
Magnitude	- $250	- $200	- $50
Control	- $100	- $50	- $50

7.2.4: fMRI analysis

Data were analyzed within the framework of the general linear model using Analysis of Functional Neuroimages (AFNI; Cox, 1996). Both individual and group-level analyses were conducted. The first four volumes in each scan series, collected before equilibrium magnetization was reached, were discarded. Motion correction was performed by registering all volumes in the EPI dataset to a volume collected shortly before the high resolution anatomical dataset was acquired. EPI datasets were spatially smoothed (isotropic 6mm Gaussian kernel) and converted into percent signal change from baseline. Regressors depicting each of the 13 response types (and sequence1) were convolved with a gamma-variate hemodynamic response function to account for the slow hemodynamic response (Cohen, 1997). The regressors were defined as follows:

a) Acquisition of response to the stimulus later associated with RR without valence change (AQRR). Percent signal change was also calculated for this stimulus following reversal (RR).
b) Acquisition of the response to the stimulus later associated with valence change but *no* response change (AQVC). Percent signal change was also calculated for this stimulus following valence change in the second phase (VC).

c) Acquisition of responding to the stimulus later associated with magnitude change without response change (AQMAG). Percent signal change was also calculated for this stimulus following a change in magnitude occurring in the second phase (VC).
d) Acquisition of the response to the control stimulus, which did not change values throughout the task (AQCON). Percent signal change was also calculated for this stimulus in the second phase (CON).
e) Acquisition errors collapsed across all stimuli in the first phase (AQERR).
f) Percent signal change for all reversal errors (RRERR) was calculated ((percent signal change to RRVC errors + percent signal change to RR errors) / 2).
g) Finally, errors to stimuli that did not reverse their reward contingencies (NRERR) was calculated ((percent signal change to VC errors + percent signal change to MAG errors + percent signal change to CON errors) / 3).

Paired t-tests were performed on a voxel-by-voxel basis across all subjects for the following contrasts: (1) RRERR versus correct reversals; (2) RRERR versus the control condition in phase 2; (3) VC condition versus the control condition in phase 2. The result was a group map of areas of differential activation. Regions of interest (ROIs) were defined at the group-level after each contrast by sampling areas that were differentially active. These ROIs were used to extract average percent signal change across each of the conditions.

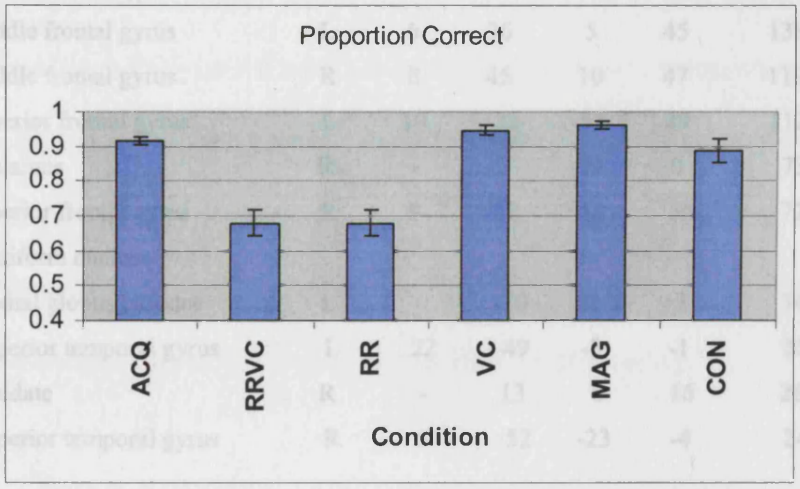
7.3: Results

7.3.1: Behavioural Results

The proportion of errors was calculated for both the acquisition and reversal phases. Repeated-measures ANOVA was conducted on the proportion of correct selections in the acquisition (ACQ) and each of the five conditions of the reversal phase (RRVC, RR, VC, MAG, and CON). This revealed a significant main effect for condition ($F(5,70) = 61.41$; $p < 0.001$). Subsequent planned contrasts revealed that participants made more errors on the two reversal conditions of the task than on any other condition (in each case, $p < 0.01$). However, the RRVC and RR conditions did not differ

significantly. Figure 7.3 displays the mean proportion correct with standard error bars for the acquisition phase and each of the phase 2 conditions

Figure 7.2: Stock exchange task behavioural data



7.3.2: fMRI Results

7.3.2.1: Response reversal errors versus correct reversals

The first analysis focused on neural responding to error feedback following a reversal in reinforcement contingencies versus activation to feedback following a correct response to reversed stimuli in the second phase. The results showed robust bilateral ventrolateral activity corresponding to BA 47 and extending into insula and BA 45 (activation is significant at $p < 0.001$, uncorrected). Significant activation was also seen in medial frontal cortex (BA 8) and OFC (BA 10). The activation is summarized in Table 7.2.

Table 7.2: Significant clusters of activation for RRERR versus correct RR

Region	L/R	BA	x	y	z	Volume
1) Medial frontal gyrus	R	8	4	24	48	12216
2) Insula/Ventral lateral	L	47/13	-31	22	3	4953
3) Inferior frontal gyrus	R	45/13	40	22	7	1992
4) Middle frontal gyrus	L	6	-36	5	45	1389
5) Middle frontal gyrus	R	8	45	10	47	1193
6) Superior frontal gyrus	L	10	-23	52	29	1128
7) Thalamus	R	-	7	-13	0	731
8) Superior frontal gyrus	R	9	32	44	34	727
9) Lentiform nucleus/ Medial globus pallidus	L	-	-10	-1	-2	346
11) Superior temporal gyrus	L	22	-49	-3	-1	261
12) Caudate	R	-	13	8	16	260
13) Superior temporal gyrus	R	-	52	-23	-4	247

L = left; R = right; BA = Brodmann's Area; X, Y, Z, coordinates are MNI;
($t = 4.321$; $p < 0.001$).

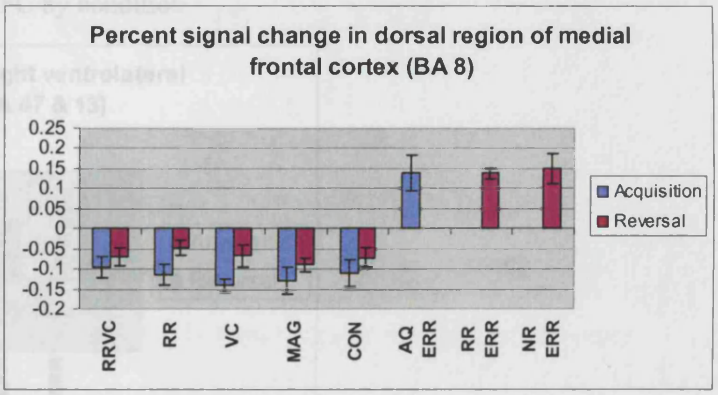
A functional mask to identify regions of interest (ROI) was generated from the contrast RRERR versus correct reversals using a statistical threshold of $p < 0.001$ (uncorrected). Using the clusters of activation, ROIs were derived that corresponded to the following regions: bilateral VLPFC (BA 47, insula, BA 13), left OFC (BA 10), and right caudate. Following the identification of the specific ROIs, the average percent signal change for all of the conditions was calculated in AFNI using "3dmaskave" (AFNI; Cox, 1996). This yielded an average percent signal change in each of the specified regions of interest for each of the 13 conditions: AQRRVC, AQRR, AQVC, AQMAG, AQCON, AQERR, RRVC, RR, VC, MAG, CON, RRERR, and NRERR. In each ROI, percent signal change relative to baseline was significantly greater in the three error conditions than in the control condition. The three error conditions did not differ

significantly from one another. The one exception was left OFC (BA 10), for which activation in response to NRERR was significantly greater than activity to AQERR. Each figure below contains an fMRI image depicting the pattern of activation elicited by the relevant contrast. In addition, each figure contains a graph plotting the percent signal change relative to baseline for each of the conditions within the region specified by the fMRI image. For each ROI, paired t-tests were conducted to determine whether activity in response to feedback following each of the three types of errors (acquisition, non-reversal, and reversal errors) differed significantly from activity in the same region to the control condition. Paired t-tests were also conducted to determine whether the percent signal change among the three error conditions differed significantly.

Figure 7.3: Activation in medial frontal cortex for RRERR versus correct RR

A) Sagittal view of ROI

B) Percent signal change in BA 8 by condition

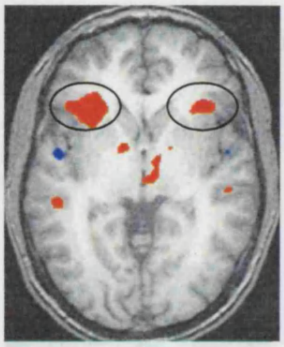


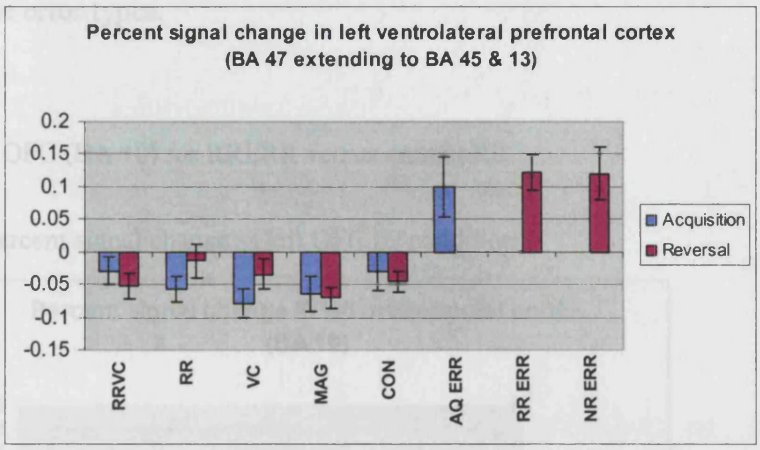
In Figure 7.3(A), a sagittal slice of activity in medial frontal cortex (BA 8) is shown. Figure 7.3(B) reveals percent signal change across conditions; all error conditions resulted in significant activation relative to correct reversals whether it was an acquisition error (AQERR), mean reversal errors for both reversal conditions (RRERR), or a non-

reversal error (NRERR; $p < 0.001$ in each case). Activation in this ROI did not differ significantly among the three error types.

Figure 7.4: Activation in VLPFC for RRERR versus correct RR

A) Axial view showing ROI B) Percent signal change in left VLPFC by condition





C) Percent signal change in right VLPFC by condition

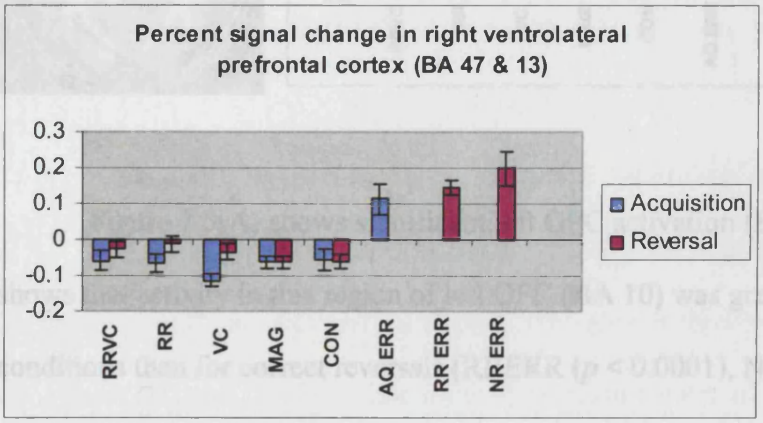


Figure 7.4(A) shows significant clusters of activation in left and right VLPFC for the contrast RRERR versus correct RR. Left VLPFC activation extends from the insula to BA 47, 45, and 13. Right VLPFC activation is focused in BA 45, 47 and 13. Figure 7.4(B) shows that for the specified region of left VLPFC, RRERR ($p < 0.001$), NRERR (p

< 0.005), and AQERR ($p < 0.05$) elicit significantly greater activation relative to the control condition. Figure 7.4(C) shows a strikingly similar pattern of results for activation in right VLPFC (including BA 47 and 13); all error conditions, RRERR ($p <$ 0.0001), NRERR ($p = 0.001$) and AQERR ($p = 0.006$) elicit significantly greater activation relative to the control condition. Activation in these ROIs did not differ significantly among the three error types.

Figure 7.5: Activation in left OFC (BA 10) for RRERR versus correct RR

A) Axial view of ROI B) Percent signal change in left OFC by condition

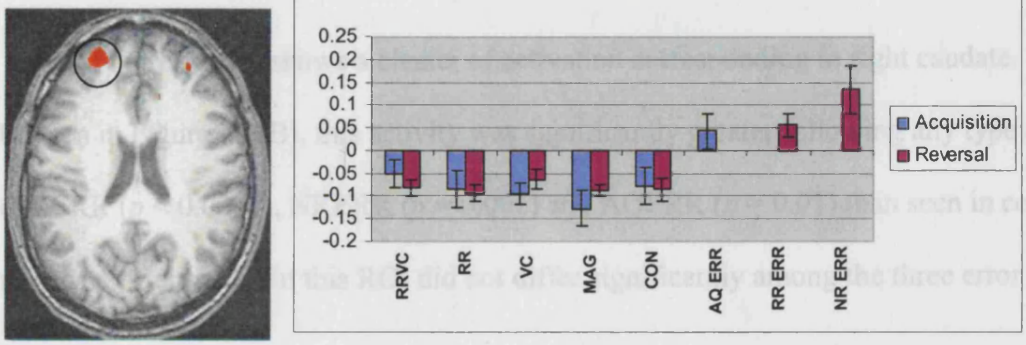
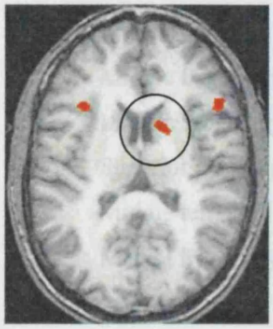


Figure 7.5(A) shows significant left OFC activation (BA 10). Figure 7.5(B) shows that activity in this region of left OFC (BA 10) was greater for each of the error conditions than for correct reversals (RRERR ($p < 0.0001$), NRERR ($p = 0.001$) and AQERR ($p = 0.02$)). Activation in this region did not distinguish between AQERR and RRERR nor NRERR and RRERR. However, significantly greater activation was seen in this region for NRERR errors over AQERR.

Figure 7.6: Activation in right caudate RRERR versus correct RR

A) Axial view of ROI B) Percent signal change in right caudate by condition



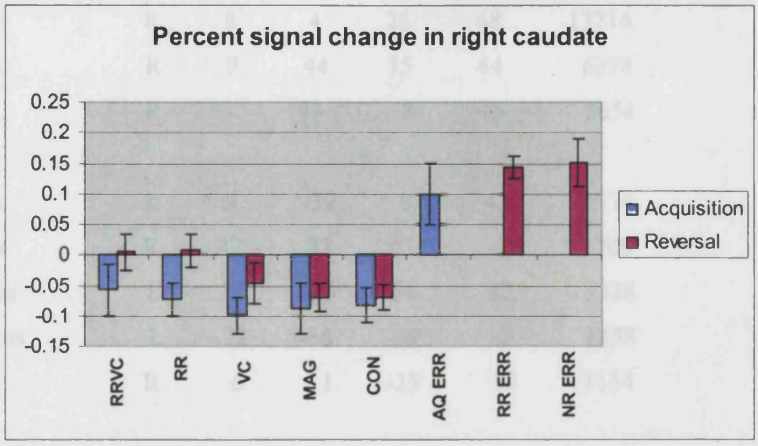


Figure 7.6(A) shows a cluster of activation corresponding to right caudate. As can be seen in Figure 7.6(B), this activity was significantly greater following any type of error (RRERR ($p < 0.0005$), NRERR ($p = 0.001$) and AQERR ($p = 0.01$) than seen in correct reversals. Activation in this ROI did not differ significantly among the three error types.

7.3.2.2: Response reversal errors versus the control condition

The second analysis investigated neural activation in response to RRERR versus the control condition. For control stimuli, the value of the reinforcer remained unchanged throughout the task. The contrast showed a similar pattern of activation as the contrast RRERR versus correct reversals. The results showed robust bilateral ventrolateral activity corresponding to BA 47 and extending into insula and BA 45. Significant activation was also seen in medial frontal cortex (BA 8) and OFC BA 10. The activation is summarised in Table 7.3.

Table 7.3: Significant activation for RRERR versus the control condition

Region	L/R	BA	x	y	z	Volume
Medial frontal gyrus	L	8	-21	9	4	18302
Medial frontal gyrus	R	8	4	28	48	13716
Middle frontal gyrus	R	9	44	15	44	6398
Caudate	R	-	13	8	15	2654
Middle frontal gyrus	L	9	-39	8	43	1974
Inferior frontal gyrus	R	47	32	21	-1	1709
Superior frontal gyrus	L	10	-23	54	27	1328
Middle temporal gyrus	L	22	-58	-34	-2	1258
Medial frontal gyrus	R	6	1	-25	59	1154
Superior temporal gyrus	R	42	63	-9	9	1002
Supramarginal gyrus	R	-	52	-47	38	860
Inferior frontal gyrus	R	11	29	36	-25	718
Middle frontal gyrus	R	10	34	55	12	431
Superior temporal gyrus	R	21	51	-24	-3	347

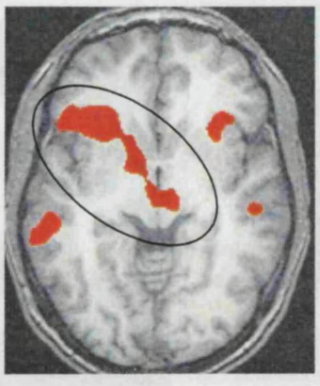
($t = 4.321$; $p < 0.001$, uncorrected).

Using the same procedure described in section 7.3.2.1, significant areas of activation generated from the contrast RRERR versus the control condition were used to form functionally defined ROIs. On this basis, ROIs were derived that corresponded to the following regions: left caudate to left VLPFC, right VLPFC, right caudate, left OFC, and right OFC. Following the identification of the specific ROIs, the average percent signal change for all of the conditions was calculated in the manner described in section 7.3.2.1. Each figure below contains an fMRI image depicting the pattern of activation elicited by the relevant contrast. In addition, each figure contains a graph plotting the percent signal change relative to baseline for each of the conditions within the regions

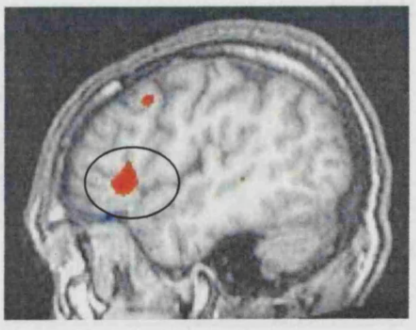
specified by the fMRI image. The same series of paired t-tests were conducted as were described in section 7.3.2.1. As with the previous contrast, paired t-tests were conducted using percent signal change in the ROI to determine whether each of the three error conditions differed significantly from the control condition and to determine whether the percent signal change generated by each of the three error conditions differed significantly from each other. For each ROI, the percent signal change in response to errors was significantly greater than the control condition. The one exception was in left and right OFC (BA 10 bilaterally).

Figure 7.7: Axial and sagittal views of activity extending from caudate to left VLPFC

A) Axial view of ROI



B) Sagittal view of left VLPFC



C) Percent signal change in area extending from thalamus to BA 47 by condition

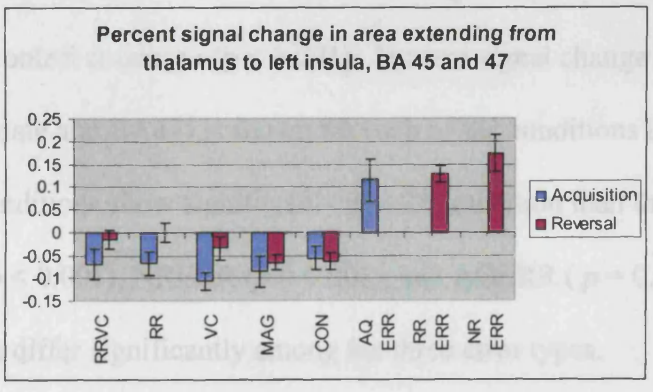


Figure 7.7(A) provides an axial view of extensive activation ranging from the thalamus to the caudate and insula to BA 47 and 45. Figure 7.7(B) provides a sagittal view showing a subset of the same activation in BA 47. In Figure 7.7(C), a bar graph illustrates the percent signal change in the ROI indicated in Figure 7.8(A and B) for each of the conditions in the task. Each of the error conditions show significantly greater activation than the control condition RRERR ($p < 0.0005$), NRERR ($p < 0.0005$, and AQERR ($p = 0.007$). Activation in this ROI did not differ significantly among the three error types.

Figure 7.8: Activation in right VLPFC for RRERR versus control

A) Axial view of ROI B) Percent signal change in right VLPFC by condition

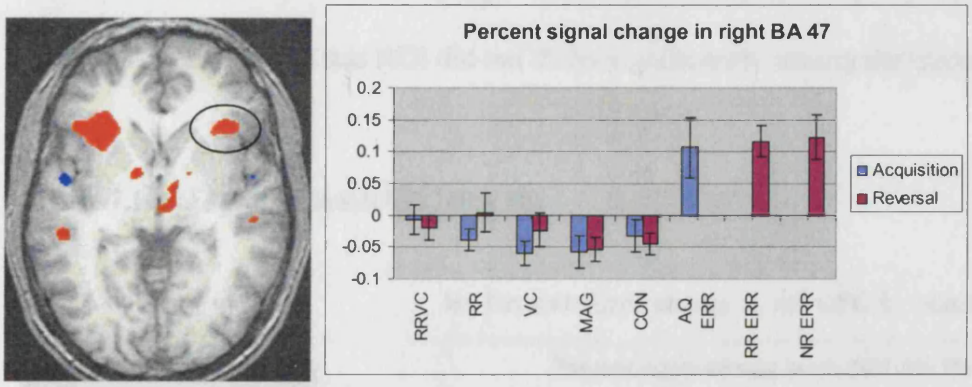
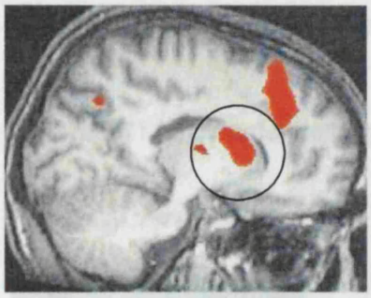


Figure 7.8(A) shows activation of the right VLPFC for the contrast RRERR versus the control condition ($p < 0.001$). Percent signal change in right lateral prefrontal cortex (caudate and BA47) is shown for each of the conditions in Figure 7.8(B). Each of the error conditions show significantly greater activation than the control condition (RRERR ($p < 0.001$), NRERR ($p = 0.001$), and AQERR ($p = 0.016$)). Activation in this ROI did not differ significantly among the three error types.

Figure 7.9: Activation in right caudate for RRERR versus control

A) Axial view of ROI



B) Percent signal change in right caudate by condition

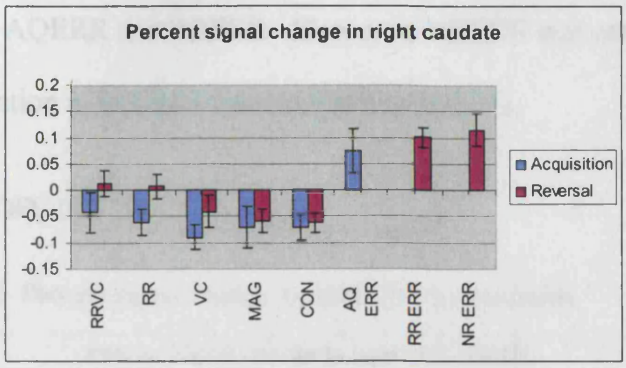
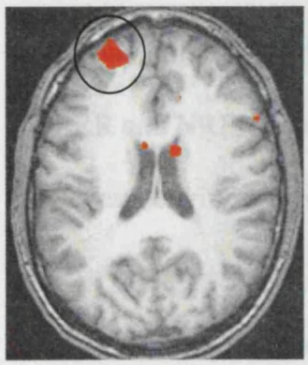


Figure 7.9(A) shows a sagittal view of right caudate activation. Figure 7.9(B) shows that each of the error conditions have significantly greater percent signal change than the control condition: RRERR ($p < 0.0005$), NRERR ($p < 0.0005$), and AQERR ($p = 0.007$). Activation in this ROI did not differ significantly among the three error types.

Figure 7.10: Activation in left OFC (BA 10)

A) Axial view of ROI



B) Percent signal change in left OFC by condition

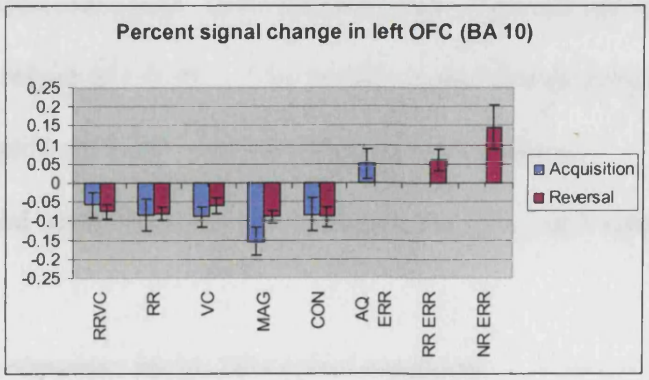
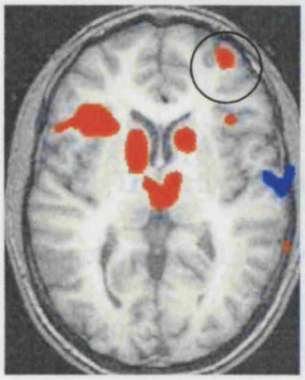


Figure 7.10(A) shows activation in left OFC (BA 10) for RRERR versus the control condition. Figure 7.10(B) illustrates percent signal change in the ROI for each of the conditions in the task. Each of the error conditions show significantly greater

activation than the control condition RRERR ($p < 0.0005$), NRERR ($p < 0.003$), and AQERR ($p = 0.014$). Paired sample t-tests revealed no significant differences in the percent signal change between AQERR and RRERR. However, NRERR was associated with significantly greater activation in this ROI than AQERR ($p < 0.05$).

Figure 7.11: Activation in right OFC (BA 10)

A) Axial view of ROI



B) Percent signal change in right OFC by condition

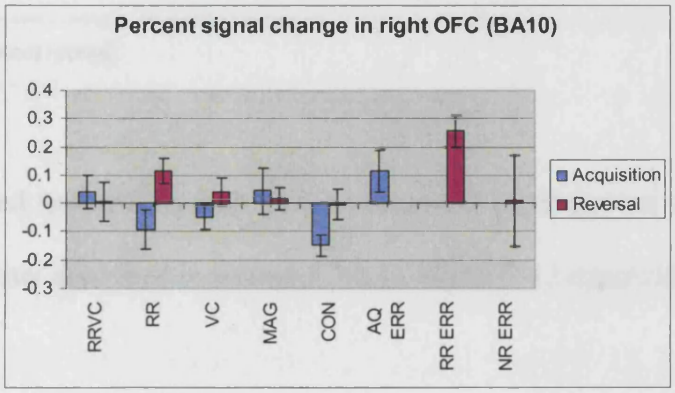


Figure 7.11(A) shows an axial view of activation in right OFC (BA 10) for RRERR versus the control condition. Figure 7.11(B) plots the percent signal change in this region for each of the conditions in the task. Only RRERR showed greater activation than the control condition in this region ($p = 0.001$). The percent signal change following AQERR and NRERR was not significantly different from the control condition. However, activity in this region did not differ significantly among the three error types.

7.3.2.3: Correct valence change responses versus the control condition

The third analysis investigated neural responding that distinguished activity to reinforcement following VC from activity following the control condition. For the VC condition, correct responses following a change in the valence of the response value occurred in the absence of demands for response change. The contrast revealed

significant activation in a region of right cingulate gyrus (BA24; ($t = 4.321$; $p < 0.001$, uncorrected). The activation is summarised in Table 7.4.

Table 7.4: Significant activation for VC versus control condition

Region	L/R	BA	x	y	z	Volume
Cingulate gyrus	R	24	19	2	38	237

($t = 4.321$; $p < 0.001$, uncorrected).

The activity generated from the contrast VC versus control condition was used as a functional ROI in the manner described in section 7.3.2.1. Figure 7.12 depicts this activation.

Figure 7.12: Activation in cingulate (BA 24) to VC versus control

A) Coronal view of ROI B) Percent signal change in left cingulate by condition

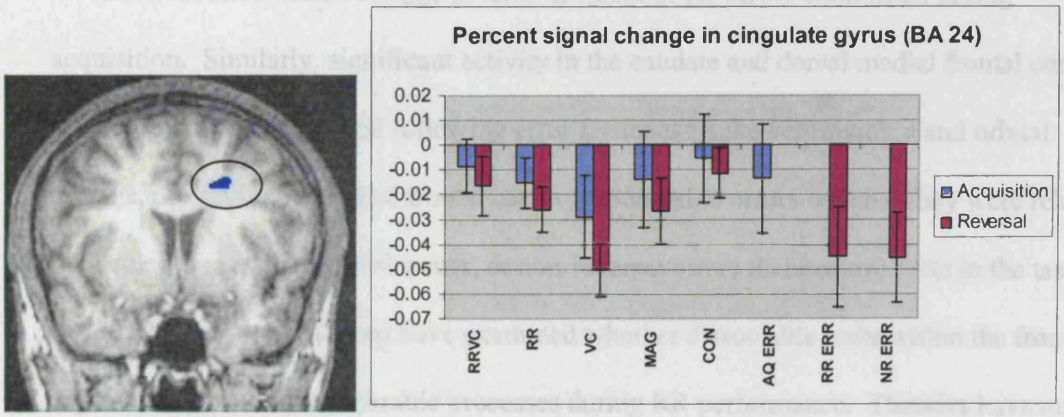


Figure 7.12(A) shows significant deactivation in the cingulate gyrus (BA 24) for the contrast VC condition relative to baseline or control ($p < 0.001$). A functional mask was generated as a result of this area of activation, and the percent signal change in each condition was examined. Figure 7.12(B) shows the percent signal change in the cingulate

gyrus (BA 24). Significant modulation in activity relative to baseline was found for the following conditions: RR (without valence change) correct ($p = 0.014$), RRERR ($p < 0.05$), all correct reversals ($p < 0.05$), and NRERR ($p < 0.05$).

7.4: Discussion

The current chapter presented fMRI data produced during the performance of a novel instrumental learning and relearning task. Participants made decisions about whether to retain or to sell stock items based on reinforcement values. In the first phase of the task, participants learned whether it was advantageous to sell or to retain each stock. In the second phase, the reinforcement value of four out of five previously correct responses unexpectedly changed. Following the change, two of the previously advantageous responses became disadvantageous; for two others, the value of the correct option changed, but the correct motor response remained the same. To provide a control condition, the value and response to the fifth stimulus remained unchanged. Significant VLPFC, caudate, and insula activity was observed following error feedback whether this feedback occurred after a change in reinforcement or for errors committed during acquisition. Similarly, significant activity in the caudate and dorsal medial frontal cortex (BA 8) region was observed following error feedback. Like ventrolateral and orbital regions, this region of medial frontal cortex responded to errors whether they were related to acquisition errors, reversal errors, or non-reversal errors that occurred late in the task.

Recent investigations have examined whether dissociable units within the frontal regions are involved in separable processes during RR performance. Theories have most frequently suggested that functional divisions exist between ventrolateral and medial regions of prefrontal cortex. O'Doherty and his colleagues (2001a) have suggested that the medial regions of prefrontal cortex may be involved in processing reward and that

lateral regions process punishment. The Stock Market Task was not set up to provide direct contrasts between positive and negative reinforcement: in this task, positively valenced outcomes could be disadvantageous and negatively valenced outcomes could be advantageous. The results of the present study are, however, *compatible* with the notion that lateral regions of prefrontal cortex are involved in processing incorrect feedback (without consideration of valence). This pattern of activation to error-related feedback was also found in medial prefrontal cortex (BA 8) and anterior regions of OFC (BA10). Thus, although the lateral/medial dichotomy proposed by O'Doherty and his colleagues (2001a), would not account for all of the error-related activation found in the Stock Market Task, it is compatible with the current results.

O'Doherty and his colleagues (2003) have recently put forth an alternative position on the functional dissociation between medial and lateral regions of OFC. The results suggest that some degree of functional overlap may exist between medial and lateral regions of frontal cortex. For example, their probabilistic RR study associated medial OFC activation with reward processing and the maintenance of a response; however, lateral regions were also active in some cases during reward processing. The authors implicate the caudolateral OFC and anterior insula in the detection of changes in contingency rather than RR per se; however, a separate area of caudolateral OFC was particularly responsive when response change was warranted. Because of the apparent overlap in function, it is difficult to generate specific predictions about lateral and caudolateral regions of OFC based on the study of O'Doherty and his colleagues study. One of the findings that was consistent throughout the study was that of medial OFC activation to response maintenance and reward processing.

The Stock Market Task data do not support these findings. For example, given O'Doherty et al.'s (2003) results, one would predict that medial regions show greater

activity following correct selections than incorrect selections. Alternatively, increased demands on a system for response maintenance may occur during the VC condition; if so, modulated activity in medial regions during the VC condition relative to the control condition would arise. The present study did not, however, find a modulation in activity in medial regions during the VC.

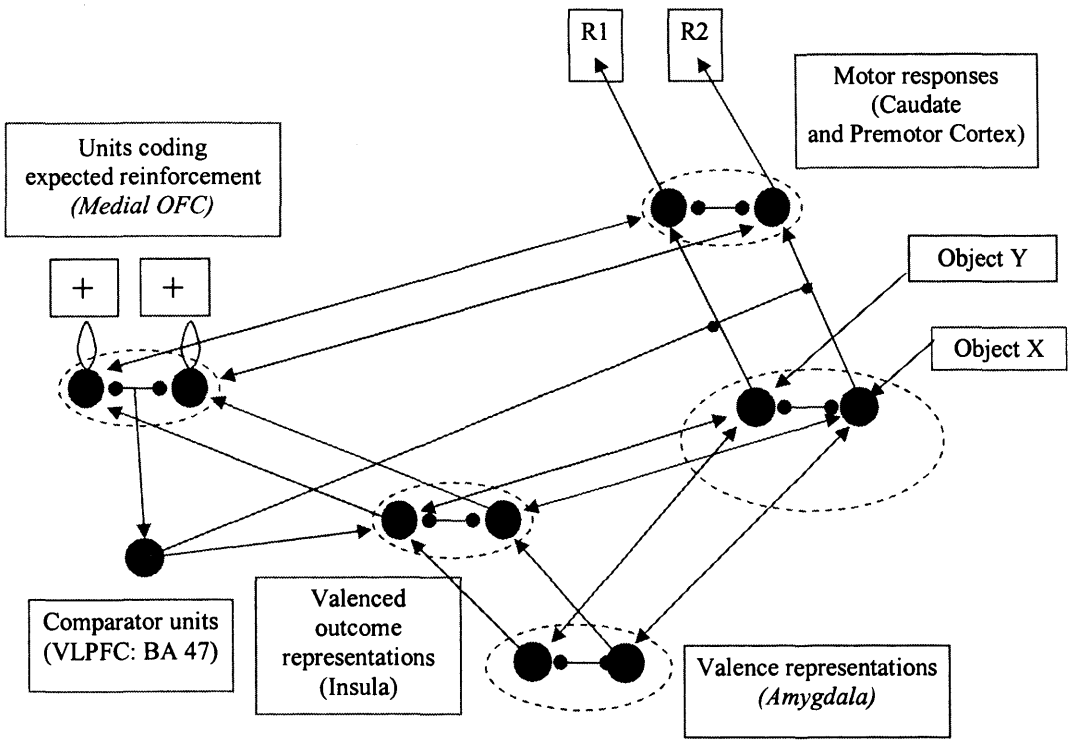
Elliott and her colleagues (1999; 2000) offer an alternate explanation. They suggest that medial regions of prefrontal cortex are involved in processing the reward value of response options when the outcome is unclear on a trial-by-trial basis; thus, the medial regions are particularly important when advantageous responding requires the integration of information accumulated over time. They suggest that medial prefrontal cortex is crucial for maintaining the "familiar" response (1999). In contrast, they state that lateral regions of frontal cortex are crucially involved when a familiar response is inhibited. However, during the Stock Market Task, activity in VLPFC was observed whenever response change was warranted, regardless of whether a prepotent response had been formed. For example, ventrolateral activation occurred to errors in the acquisition phase before a prepotent response would have been developed. Similarly, significant activation in ventrolateral regions occurred to non-reversal errors in the second phase. Thus, activation was observed in the present study whenever an errant choice was made, even if the errant choice was not a practised one. As a consequence, this finding does not support conceptualisations that emphasize a role for the lateral prefrontal cortex in response inhibition or the suppression of a prepotent response.

The IES model (Blair, 2004) is an alternative explanation of VL/OFC function that emphasizes the importance of representations of *expectations* of reinforcement. The representation of response outcome expectations is implicit in many influential models of emotional processing. For example, models of emotional experience (e.g., Mandler,

1984; Oatley & Johnson-Laird, 1987), anxiety (e.g., Gray, 1982), stimulus-reinforcement learning and relearning (e.g., Grossberg, 1982; Rolls, 1990), frustration (Amsel, 1992), and latent inhibition (Schmajuk, Lam, & Gray, 1996) include representations of response outcome expectations. Representation of response outcome can also be applied to OFC functioning. The IES model incorporates the notion of response outcome encoding and suggests that separate systems exist for the encoding of expectations, the detection of violations in these expectations, and for initiating changes in responding (Blair, 2004).

The Stock Market Task was designed to investigate elements of RR that are directly relevant to key components of the IES model (Blair, 2004). The IES model, depicted schematically in Figure 7.13, suggests that regions of VLPFC play a crucial role in response change within a larger circuit that includes medial regions of prefrontal cortex, the amygdala, insula, caudate, and premotor cortex. The amygdala contains units dedicated to the representation of reinforcer valence. These units are activated by aversive and appetitive reinforcers. Units within the insula represent outcome of the available response options, which acquire positive or negative valence based on input from valence representation in the amygdala. Thus, the representations of these outcomes becomes rewarding or aversive. Valence information is conveyed to units within the medial OFC which code the expected reinforcement value of motor responses represented and initiated by the caudate and premotor cortex. Functionally, the model suggests that units within BA 47 detect mismatches between expectations of valence reinforcement and actual outcomes. Thus, damage to the insula, ventrolateral, or medial regions of prefrontal cortex will all impact RR.

Figure 7.13: Suggested mechanism for RR (IES model; Blair, 2004)



With regard to the Stock Market Task, the IES model offers the clear prediction that regions of VLPFC will be engaged whenever mismatches exist between the optimal amount of reinforcement available, and the actual reinforcement received. The results supported this prediction. Activation in VL/OFC occurred whether or not the error was committed during acquisition or following a reversal. In this study, activity in the medial prefrontal cortex did not distinguish between changes in reinforcement that did not compel response change from the control condition. Although a region of dorsal cingulate was sensitive to changes in valence, it also showed modulation relative to the control condition for reversal and non-reversal errors.

In the present study, differential activation in the valence and magnitude change conditions was not apparent. This is surprising given the clear evidence for a role of the OFC in processing both primary and secondary reinforcement value (Thorpe et al., 1983; Critchely & Rolls, 2000). One possibility is that changes in the value of responses in the current task were not sufficiently salient. Participants may have been attending to the response-relevant information at the expense of the more subtle performance-related information. This possibility is currently being explored in an experiment in which the salience of the reinforcement change is increased without affecting behavioural choice.

7.5: Conclusion

The results of the study presented here supports theories of VLPFC (BA 47) and more anterior regions of OFC (BA 10) in signalling an incorrect response whether or not the feedback violates an established stimulus-response association. This suggests that ventrolateral regions are not simply involved in inhibiting a prepotent response as has been suggested. The hypothesis that medial frontal activation would be associated with all changes in the value of reinforcement regardless of demands for response change was not supported by the present study. This may be due to reduced salience of reinforcement values that were not motivationally significant. The data did support the prediction that VLPFC would be active during RR. Activation was seen in these regions to all errors, regardless of whether they were following a violation in RR. The data generated by this experiment challenge theories that suggest that the primary role of the VL/OFC is in the inhibition of a prepotent response.

Chapter 8

Summary, Conclusions, and Future Directions

8.1: Introduction

This chapter summarizes the implications of this thesis and suggests directions for future work. The conclusions drawn concern the impact of emotion on attention, dissociable forms of instrumental learning, decision-making, and behavioural regulation. Cognitive processes were assessed using standard and novel experimental tasks in both healthy and patient populations. The cognitive functions discussed are specifically related to disorders of aggression such as psychopathy and acquired sociopathy. The relevance of amygdala and OFC dysfunction in psychopathic disorders is reviewed.

8.2: Summary and Conclusions

8.2.1: Reviewing OFC function, amygdala function, and psychopathy

This thesis began with a discussion of the construct of psychopathy and the impetus behind the theories and neurocognitive investigations of the disorder. This discussion was followed by sections concerning the basic anatomy and function of the OFC. The aberrant social behaviour that can follow OFC damage was also described. The laboratory work reviewed implicates the OFC in expression recognition, decision-making, and response change. Three influential theories of OFC function were discussed: the SRR theory (Blair & Cipolotti, 2000), the somatic marker hypothesis (Damasio, 1994), and a RR theory (Rolls, 1999). An illustration of how these theories can be merged under the Integrated Emotions System model was provided (Blair, 2004). The clinical correlates of amygdala dysfunction and anatomy of the amygdala were discussed in the second half of the chapter, followed by a description of the involvement of the

amygdala in emotional expression recognition, conditioning, *some* forms of instrumental learning, and the emotional modulation of attention. The chapter concludes with a discussion of how the thesis addresses its chief aim of characterising OFC and amygdala dysfunction in psychopathy.

8.2.2: The emotional modulation of attention

A novel experiment investigating the modulation of motor behaviour by emotion in forensic samples with and without psychopathy was presented in Chapter 2. The introduction discussed evidence showing that healthy emotional responding can lead to an impairment of operant behaviour given the appropriate conditions. This was followed by a description of two mechanisms by which emotion can disrupt behaviour: (1) through the generation of an incompatible response, "freezing," generated by the basic threat circuitry, or (2) by "capturing" attention. In line with predictions, individuals with psychopathy showed reduced interference of ongoing behaviour in the presence of positive and negative images, but not for trials associated with neutral images. The disrupted behaviour during the positive images suggested that the mechanism of disruption involved in the Emotional Interrupt Task was through the depletion of attentional resources rather than through the activation of the basic threat response system. Specifically, conclusions drawn on the basis of this data suggest that although the comparison individuals showed reduced allocation of attentional resources to the operant task during positive and negative presentations, the individuals with psychopathy did not exhibit this modulation of attention.

8.2.3: Stimulus reinforcement learning

Chapter 3 begins with a description of two fundamental forms of instrumental learning conceptualised by Baxter and Murray (2002) as stimulus-reinforcement learning and stimulus-response learning. The two classic tests of stimulus-reinforcement and stimulus-response learning discussed are thought to rely on the integrity of dissociable neural substrates. Thus, while stimulus-reinforcement learning is impaired following amygdala damage, object discrimination learning remains intact (Baxter & Murray, 2002). It was hypothesized that individuals with psychopathy would show difficulties only on instrumental learning tasks that required the formation of stimulus-reinforcement associations. The results of a stimulus-reinforcement association task involving groups comprised of individuals with psychopathy, individuals with sub-threshold psychopathy, and comparison individuals were presented. In line with predictions, a correlation was found between the level of psychopathy and the number of errors in the acquisition of the instrumental response: psychopathic individuals made significantly more errors relative to the comparison group in the learning phase. Two reversal phases followed. In the second reversal phase, individuals with psychopathy and sub-threshold psychopathy made significantly more errors than the comparison group. As expected, psychopathy was correlated with the number of errors during acquisition and the second reversal phase. The results were in line with suggestions that individuals with psychopathy present with an impaired ability to form stimulus-reinforcement associations and an impaired reversal of such associations once they have been formed.

8.2.4: Reversal learning and decision-making

The primary objective of Chapter 4 was to investigate further the presence and nature of the proposed reversal learning deficit exhibited by individuals with psychopathy. Two experiments were conducted. In the first, a gambling task was used in which participants chose between four response options of different reinforcement contingencies. Two responses were high-risk and ultimately disadvantageous, and two were low-risk and ultimately rewarding. The comparison group gradually showed a preference for the low-risk decks. In contrast, individuals with psychopathy continued to sample from the high-risk decks at the same rate throughout the task. The results are compatible with either an OFC or an amygdala dysfunction explanation of psychopathy; patients with lesions involving either the OFC or the amygdala are impaired on this task (Bechara et al., 1999). Chapter 4 included results from a second experiment, the ID/ED task, which is a multi-component measure of instrumental learning, attentional set-shifting, and RR. The task involves a form of instrumental learning (object discrimination learning) that is not thought to rely on the amygdala (Baxter & Murray, 2002). The attentional set-shifting component of the task is thought to be sensitive to dorsolateral regions of frontal cortex and the reversal component to OFC regions (Dias et al., 1996). In line with predictions generated from the idea of a specific VL/OFC dysfunction theory of psychopathy, individuals with the disorder showed intact object discrimination learning and attentional set-shifting, but impaired RR.

8.2.5: Conditional learning

Chapter 5 presented an experimental investigation of conditional learning and conditional reversals. Unlike the experiment presented in Chapter 3, this task is thought

to measure stimulus-response associations, and therefore is not reliant on the amygdala (Baxter & Murray, 2002). Conditional learning involves selecting from available responses in the context of a reference stimulus. In making conditional discriminations, participants do not approach or avoid stimuli that are intrinsically positive or negative, but rather, determine the appropriate response in a given stimulus' presence. The distinction between stimulus-response and stimulus-reinforcement associations allows one to reliably predict which instrumental learning tasks depend on the amygdala and which do not. As a consequence, it provides an excellent means with which to explore whether the apparent instrumental learning deficit observed in individuals with psychopathy is due to a general emotional learning difficulty, or whether, as the amygdala dysfunction hypothesis of psychopathy would suggest, it is specific to stimulus-reinforcement associations. As predicted by both the amygdala and OFC dysfunction theories of psychopathy, participants with the disorder showed intact acquisition and maintenance of stimulus-response associations and impaired reversal of those associations. The results provide further support for a specific, rather than general, instrumental learning deficit in individuals with psychopathy.

8.2.6: Acquired sociopathy versus developmental psychopathy

Acquired sociopathy, a disorder that can arise following insult to the OFC, has been suggested as a model for developmental psychopathy (Damasio, 1994; Anderson et al., 1997; Anderson et al., 1999). Chapter 6 presents a patient alleged to have acquired sociopathy. Standard and experimental tasks were conducted to characterise his impairment and to compare and contrast his performance with a healthy community sample as well as psychopathic and non-psychopathic forensic comparison groups. As reviewed in previous chapters, considerable evidence suggests that OFC dysfunction may

contribute to at least some of the symptoms associated with psychopathy. However, given the hypothesis that amygdala dysfunction is the primary deficit involved in the disorder, it was predicted that patient C.L. would show a dissociable pattern of impairment relative to individuals with psychopathy. The results revealed that the psychopathic group demonstrates a strikingly different pattern of performance than C.L. For example, C.L. showed pronounced SRR deficits: he had facial expression recognition impairments (particularly anger and disgust) and had difficulties identifying situations that would result in anger in others. He also showed impaired Iowa Gambling Task performance, but surprisingly, he did not exhibit RR deficits on the ID/ED Task. In contrast, individuals with developmental psychopathy did not show evidence of SRR deficits. Like C.L., individuals with developmental psychopathy showed impaired decision-making on the Iowa Gambling Task, but, unlike C.L., they also showed RR deficits on the ID/ED Task. Theory of mind deficits were not evident in any of the groups. One of the most important findings of this chapter, however, was the double dissociation seen between C.L. and the psychopathic group on instrumental learning tasks. C.L. showed intact performance on the passive avoidance test, but pronounced impairment on the conditional learning task. Conversely, individuals with psychopathy showed gross impairment on the passive avoidance task and intact performance on the conditional learning task. C.L.'s performance profile is probably reflective of the extent of his frontal lesion and sparing of the amygdala. The performance profile of the individuals with psychopathy is predicted by conceptualisations that suggest the presence of OFC and amygdala dysfunction, but with sparing of cognitive function ascribed to the dorsolateral prefrontal cortex.

8.2.7: Response change and value encoding

Substantial evidence exists to suggest that regions of OFC and VLPFC play an integral role in response change and decision-making. In Chapter 7, evidence was reviewed that suggested that regions of OFC play a role in processing primary and abstract reinforcement. Data was also reviewed showing that damage to this region resulted in impaired RR. Theories and studies of OFC function either look at these processes in isolation or collapse them together so that it is difficult to determine whether dissociable regions within OFC or VLPFC are involved in stimulus value encoding and response change. A novel Stock Market Task that attempts to disentangle stimulus-value encoding from response change was presented. The study showed a striking and consistent pattern of activation in VLPFC (BA 47) and orbital (BA 10) regions of prefrontal cortex whenever an error was made, regardless of whether it occurred during acquisition, to a change in reinforcement level, or to an errant response that occurred long after acquisition. In short, the signal appeared whenever a response was made that ought not to be repeated on the subsequent trial. One of the central aims of the experiment was to determine whether dissociable regions within the prefrontal cortex were involved in detecting changes in the level of reinforcement independent of response change. Significant activation was not, however, seen in the contrast valence change versus the control condition. This is likely to reflect a failure to encode the change in valence rather than being indicative of a common neural substrate. Rather than attending to the actual magnitude of reinforcement, participants may have been comparing the level of reinforcement that they received to ensure that they were making the advantageous decision. The absolute value of the reinforcement each time would therefore not be important.

8.3 Relevance to the Clinical Presentation of Psychopathy

Neurocognitive deficits that are associated with a particular disorder may be characteristic of the disorder without having relevance to the clinical presentation. The neuropsychological tests presented in this thesis were designed or selected to investigate skills associated with real-world functioning such as decision-making, responding on the basis of reward and punishments, sensitivity to social cues, and response change. In most cases, the cognitive skills investigated can be considered a more basic, measurable marker of a more complex neurocognitive deficit that contributes to the clinical presentation of the disorder being explored. For example, psychopathy is characterised by elevated rates of instrumental and reactive aggression. It is thought that the circuitry involved in these two types of aggression are dissociable, and that both are impaired in individuals with psychopathy (Blair, 2003b; 2004). Deficits in response reversal do not, in isolation, result in persistent aggressive and otherwise maladaptive social behaviour. However, the relationship of reversal deficits to aggression is clarified by considering that response reversal is one component of a system that is involved in encoding the motivational value of response options and updating the values when they change. Furthermore, given that frustration, the precursor to reactive aggression (Berkowitz, 1993), results when an expected reward is not obtained, a failure in the system will leave someone at risk for heightened frustration, and consequently, increased rates of reactive aggression. Similarly, insensitivity to distress cues is thought to be a risk factor for instrumental aggression (Blair, 2003a). Difficulty representing distress in others as an aversive stimulus may somewhat increase the risk of instrumental aggression (Blair et al., 1997; Blair, 2003b). However, sensitivity to distress appears to be mediated by a neurocognitive system (in this case involving the amygdala) that is involved in many

forms of emotional learning and responding such as fear conditioning and other forms of emotional learning (LeDoux, 1998). In both examples, reversal learning and expression recognition, the skills explored are a measurable property of a larger neurocognitive system that together with environmental factors, contribute to the clinical presentation of psychopathy.

8.4 Strengths and Weaknesses

8.4.1: Strengths

One of the main strengths of the present thesis was the incorporation of a diverse range of research techniques. In developing theories with respect to psychopathy and the neurocognitive basis of instrumental learning, data from several levels were included ranging from animal research, developmental disorders, acquired lesions in humans, and functional imaging work all of which guided the experimental design and interpretation of existing results. Similarly, the experiments presented in this thesis combine established and novel neuropsychological as well as functional imaging techniques to address research questions. Like the techniques, the samples involved were varied. The patient groups included developmental psychopaths, a patient with acquired sociopathy, incarcerated controls, and healthy individuals from the community. The inclusion of these groups allowed for a rare experimental comparison between a case of acquired sociopathy and individuals with developmental psychopathy.

An additional advantage to utilizing data from a variety of domains is that it enables the generation of highly constrained predictions. The studies examined the performance of individuals with psychopathy on *dissociable* forms of instrumental learning for the first time. Based on data in the animal literature, specific predictions

could be made about performance despite the surface similarities between the different instrumental behaviours. As a result, this thesis provides what may be the first evidence that psychopathy is characterized by a specific instrumental learning deficit rather than a global one. Psychopathy is characterised by an impairment in making stimulus-reinforcement associations. Stimulus-reinforcement associations entail attaching a positive or negative "affective tag" to a stimulus, which is thought to require the amygdala (Baxter & Murray, 2002; Everitt et al., 2003). In contrast, these experiments show no evidence for dysfunction in stimulus-response associations in individuals with psychopathy. Stimulus-response associations do not require intact amygdala functioning (Baxter & Murray, 2002).

Finally, the use of novel experimental techniques proved useful. For example, Chapter 2 presented a novel paradigm that allowed for an experimental manipulation that demonstrates that a neurocognitive deficit in psychopathic individuals could lead to enhanced performance relative to comparison individuals. In Chapter 7, a novel task was constructed to truly separate changes in valence and magnitude from the existing confounds of response change in a response reversal task.

8.4.2: Weaknesses

Minor adjustments to the methodology of some of the experiments will improve their utility for future studies. Chapter 2 presented an experiment in which participants were asked to attend to behaviourally irrelevant picture stimuli while responding to shapes. By removing any instruction to attend to the image, it would be clearer that a failure to modulate responding to the emotional material is due to insensitivity to the *incidental* emotional stimuli rather than failure to comply with a component of the instructions.

Chapters 3, 4, and 5 presented data on dissociable forms of instrumental learning. Although providing valuable information about the dissociable forms of instrumental learning, the data could be more persuasive if the different tasks were matched for difficulty. For example, the fact that both groups perform at ceiling for the learning component of the ID/ED limits conclusions that can be drawn about the absence of an object discrimination impairment.

The ability to generalize the results contained in this thesis is limited for some of the studies due to small sample size. This is particularly clear in Chapter 6, in which a single case study with modest control sample size was presented. Patients with a focal lesion are relatively rare, which makes even a single case study of value. However, the results would be more persuasive with a larger healthy control group and a forensic sample that was better matched on age and education.

Chapter 7, although representing a novel examination of the different components of response reversal, did not detect substantial activation associated with changes in the magnitude of reinforcement. This may be due to insufficient power related to a relatively small sample size (n=12). Alternatively, the value change may not have been salient enough for participants to encode.

Future work aimed at addressing these weaknesses and building on the strengths of the experiments discussed in this thesis is described in the following sections.

8.5 Future Directions

The empirical work in this thesis has highlighted a number of possible avenues for further research. In the course of investigating neurocognitive deficits associated with psychopathy, more basic questions about the underlying cognitive processes under

examination have arisen. The final sections of this thesis will discuss these possibilities for further research.

8.5.1: The modulation of attention and emotion

The Emotional Interrupt Task presented in Chapter 2 attempted to investigate the impact of emotional material on ongoing behaviour. This task is currently being implemented with special patient groups, including children with psychopathic tendencies and bipolar disorder, to test the generalisability of the result. One can predict that disorders that are believed to be associated with increased activation of threat circuitry, such as post-traumatic stress disorder (PTSD; Gilboa et al., 2004; Nutt & Malizia, 2004), will also be associated with increased disruption on the Emotional Interrupt Task. This hypothesis is currently being investigated in a sample of women with PTSD.

This work inspired additional questions about the basic processes involved in the modulation of attention. The nature of the interaction between attention and emotion is currently debated. One perspective suggests that emotional material automatically gains access to attentional resources regardless of the extent to which attentional resources are depleted (Vuilleumier et al., 2001). An alternative view argues that emotional material is not automatically processed, but rather, requires that attentional resources be available (Pessoa et al., 2002). In support of this position, Pessoa and his colleagues (2002) showed that emotional material is not processed by participants who are performing a cognitive task that places high demands on attention; brain regions that are consistently activated in the presence of emotional faces such as the fusiform gyrus (Kanwisher, Stanley, & Harris, 2000) and amygdala (Morris et al., 1996; Whalen et al., 2001) did so only when participants did not engage in the attentionally demanding task. This was true even when emotional stimuli were placed centrally at fixation and the stimuli relevant to

the cognitive task were presented in the periphery. The two studies, however, used different cognitive tasks that varied substantially in terms of spatial organization, thereby making it difficult to evaluate the validity of the two opposing results.

This problem has inspired a new fMRI project, already underway, investigating the impact of emotional information under both high and low attentional load conditions. The task involves the rapid presentation (250ms) of compound stimuli consisting of a semi-transparent word superimposed on a fearful or neutral facial expression. The task has three conditions. In the control condition, participants are asked simply to judge the gender of the face. In the low attentional load condition, participants judge whether the superimposed words are presented in upper or lower case. In the high attentional load condition, participants are asked to determine the number of syllables contained in the superimposed word. The task is designed to address the continuing question of whether the amygdala responds automatically to emotional stimuli or does so only when attentional resources are available. It has an advantage over previous designs because it incorporates a powerful parametric design within one task and uses identical stimuli across conditions. If, as Vuilleumier et al. (2001) have suggested, the amygdala processes emotional stimuli automatically, equivalent amygdala activation should be seen in both the high and low attentional load conditions for fearful, but not neutral faces. If, as Pessoa and his colleagues suggest (2002), all processing requires the availability of attentional resources, the level of amygdala activation in the high load condition should be less than in the low attentional load condition. This study is designed to address this question.

8.5.2: Dissociable forms of instrumental learning

This thesis examined different forms of instrumental learning that, although similar, are dissociable at the neural level. Much of the available work concerning dissociating these forms of instrumental learning has been conducted with animals. Future work incorporating imaging technology with human subjects to illustrate further the neural substrates involved in different forms of instrumental learning would be valuable. Chapter 4 presented data from the ID/ED Task, which features object discriminations and reversals. Object discrimination is thought to be a form of instrumental learning that does not rely on the amygdala. In line with this idea, individuals with psychopathy did not show impaired performance on this task. However, the simplicity of the object discriminations in the ID/ED Task raises questions about ceiling effects. It is unclear whether the task is sensitive enough to detect existing object discrimination impairments. As an extension of the work presented in Chapters 3, 4 and 5, I have recently designed a new computer-based object discrimination task involving six stimulus pairs. Through piloting with healthy volunteers, this new object discrimination task has been matched in difficulty with the Tokens Task. If the amygdala hypothesis of psychopathy is correct, patients with disorders that are characterised by reduced amygdala functioning, such as psychopathy, should show intact performance on this new instrumental learning task, but impaired learning on the Tokens Task. In contrast, patient groups that are thought to be hyper-responsive to stimulus-reinforcement associations should show altered performance on stimulus-reinforcement learning tasks, but intact functioning on stimulus-response learning tasks such as object discrimination.

8.5.3: Acquired sociopathy and other aggressive disorders

Chapter 6 reviewed evidence concerning the link between aggressive disorders and dysfunction involving the OFC and amygdala. Clinical reports of symptoms that

follow damage to the frontal lobes have become more precise in identifying sub-regions within the frontal lobes that contribute to the aberrant behaviour. However, studies involving larger populations of individuals with damage to these regions can increase the specificity of the link between neurocognitive deficits and social disturbance. For example, it is not clear whether RR deficits in themselves contribute to aggression or whether such deficits are symptomatic of the underlying disease. Rolls and his colleagues (1994) showed a correlation between reversal errors and behavioural problems as measured by a caregiver's report. However, further studies that examine the correlation between RR errors and aggression are needed. For example, similar work relating extent of aggressive outbursts and RR or SRR impairments in other violent disorders would help clarify the relationship between this cognitive deficit and violent behaviour. Based on the theories presented in this thesis, the strength and replicability of either RR or SRR deficits in any patient population should be related to the degree of reactive aggression exhibited in this population.

As was discussed in the first chapter, the clinical picture of patients with lesions involving the amygdala is inconsistent. This may be due in part to the paucity of cases: circumscribed amygdala lesions are comparatively rare relative to lesions involving the OFC (most are gleaned from patients with temporal lobectomies). Research involving chimps suggests that lesions acquired later in life produce subtly different behavioural profiles than those acquired in infancy (Emery, Capitanio et al. 2001; Prather, Lavenex et al. 2001). More detailed studies of patients who develop lesions at different stages of development would be very informative. For example, research suggests that damage to the amygdala can lead to either increased (van Elst, Woermann et al., 2000) or decreased (Ramamurthi, 1988) aggression. These two findings are difficult to reconcile. However, Blair (2004) has suggested that dysfunction within the amygdala can increase the

probability of aggression by reducing the individual's sensitivity to distress cues in others or by disrupting the effect of appetitive stimuli, both of which may have an inhibitory effect on threat circuitry. He has also suggested that the incidence of aggression can be reduced through amygdala lesions by decreasing the patient's sensitivity to learnt threat. At present, however, the parameters that determine which of these effects occur is not known. Further research investigating patients with lesions involving the amygdala would help clarify the relevant determinants of aggression.

8.5.4: Response reversal versus valence change

Chapter 7 presented a novel instrumental learning task that dissociates between individual elements of RR traditionally treated as a unitary process. The experiment is the first in a series of studies planned to investigate the neural regions involved in separable components of RR. The task successfully distinguished between reversal errors and non-reversal errors to a degree not possible in previous reversal paradigms. The design was also novel in that it separated changes in reinforcement value that were sufficient to motivate behavioural change from changes that were not. The pattern of activation to changes in value of a response, however, showed great overlap with the control condition. This may reflect reduced salience of changes in reinforcement value that do not have behavioural significance. Further investigation will increase the salience of the magnitude and valence change conditions perhaps through the inclusion of a test phase. Plans are underway to implement these changes.

8.6: Summary

In this thesis, behavioural measures were conducted to explore the relationship between psychopathy and performance on measures associated with OFC and amygdala

functioning. Consistent with theories stressing the presence of OFC dysfunction in psychopathy, patients with psychopathy showed deficits in the reversal of both object discriminations and conditional associations, as well as on a decision-making task. Despite suggestions in the literature that equate psychopathy with acquired sociopathy, this thesis presented both theoretical and empirical reasons why an OFC dysfunction account in isolation is insufficient to explain the characteristics associated with the disorder. C.L., a patient with "acquired sociopathy" following lesions of the OFC, showed impaired conditional learning, but intact passive avoidance learning. He also showed gross SRR impairments, but intact reversal learning. Conversely, individuals with psychopathy showed impaired passive avoidance learning, but normal ability to make conditional discriminations. Furthermore, individuals with developmental psychopathy showed impaired reversal learning, but intact SRR. The data presented also bolster theories of psychopathy that emphasize amygdala dysfunction; individuals with psychopathy showed impaired modulation of behaviour by attention, impaired decision-making, impaired expression recognition, and impaired stimulus reinforcement learning. Related to these findings, evidence was provided for a dissociation of instrumental learning into multiple forms: stimulus-reinforcement and stimulus-response learning. Individuals with psychopathy were impaired on the former, but not the latter form. An imaging study investigating RR was conducted. This study was unusual in that its design allowed for a genuine separation between reversal processes that have been traditionally confounded: changes in the value of reinforcement that are sufficient to motivate response change and changes in the value of reinforcement that are insufficient to induce behavioural change. The present chapter has identified some of the key questions raised by this thesis, and in doing so, suggested directions for future research.

References

Adolphs, R., & Tranel, D. (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia, 37*(11), 1285-1292.

Adolphs, R., Tranel, D., & Damasio, H. (2001). Emotion recognition from faces and prosody following temporal lobectomy. *Neuropsychology, 15*(3), 396-404.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669-672.

Adolphs, R., Tranel, D., Young, A.W., Calder, A.J., Phelps, E.A., Anderson, A.K., Lee, G.P., & Damasio, A.R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsycholgia, 37*, 1111-1117.

Aggleton, J.P. (2000). The amygdala--what's happened in the last decade? In J.P. Aggleton (Ed.), *The Amygdala* (2nd ed., pp. 1-30). Oxford: Oxford University Press.

Alexander, M., & Perachio, A.A. (1973). The influence of target sex and dominance on evoked attack in Rhesus monkeys. *American Journal of Physical Anthropology, 38*(2), 543-547.

Amaral, D.G., Price, J.L., Pitkanen, A., & Carmichael, S.T. (1992). Anatomical organization of the primate amygdaloid complex. In J.P. Aggleton (Ed.), *The Amygdala: Neurobiological aspects of Emotion, Memory, and Mental Dysfunction* (pp. 1-66). New York: A John Wiley-Liss & Sons, Inc.

Ambrogi Lorenzini, C.G., Bucherelli, C., Giachetti, A., Mugnai, L., & Tassoni, G. (1991). Effects of nucleus basolateral amygdalae neurotoxic lesions on aversive conditioning in the rat. *Physiology and Behavior, 49*(4), 765-770.

Anderson, A.K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J.D.E. (2003). Neural correlates of the automatic processing of threat facial signals. *The Journal of Neuroscience, 23*(13), 5627-5633.

Anderson, A.K., & Phelps, E.A. (1998). Bilateral amygdala damage impairs evaluation of facial but not vocal expressions of fear. *Cognitive Neuroscience Meeting Abstract Program, 109.*

Anderson, A.K., & Phelps, E.A. (1998). Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport, 9*(16), 3607-3613.

Anderson, A.K., & Phelps, E.A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411*, 305-309.

Anderson, S.W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A.R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience, 2*(11), 1032-1037.

Anderson, S.W., Damasio, H., Tranel, D., & Damasio, A.R. (1997). Long-Term Sequalae of Prefrontal Cortex Damage Acquired in Early Childhood. *Developmental Neuropsychology, 18*(3), 281-296.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29*(5): 407-18.

Baron-Cohen, S., Ring, H., Moriarty, J. Shmitz, P., Costa, D., & Ell, P. (1994). Recognition of mental state terms: A clinical study of autism, and a functional neuroimaging study of normal adults. *British Journal of Psychiatry 165*, 640-649.

Baron-Cohen, S., Ring, H.A., Bullmore, E.T., Wheelwright, S., Ashwin, C., & Williams, S.C. (2000). The amygdala theory of autism. *Neuroscience Biobehavior Review, 24*(3), 355-364.

Baron-Cohen, S., Wheelwright, S., Hill, J., Rate, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2): 241-51.

Barrash, J., Tranel, D., & Anderson, S.W. (2000). Acquired personality disturbances associated with bilateral damage to the vetromedial prefrontal region. *Developmental Neuropsychology, 18*(3), 355-381.

Baxter, M.G., & Murray, E.A. (2002). The amygdala and reward. *Nature Reviews Neuroscience, 3*(7), 563-573.

Bechara, A., Damasio, A.R., Damasio, H., & Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*, 7-15.

Bechara, A., Damasio, H., Damasio, A.R., & Lee, G.P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience, 19*, 5473-5481.

Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S.W., & Nathan, P.E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia, 39*(4), 376-389.

Bechara, A., D. Tranel, & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123,* 2189-2202.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A.R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science, 269*, 1115-1118.

Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control.* Philadelphia: Temple University Press.

Berlin, H.A., Rolls, E.T., & Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain, 127*, 1108-1126.

Berthoz, S., Armony, J.L., Blair, R.J., & Dolan, R.J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain, 125*, 1696-1708.

Blair, R.J.R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition, 57*, 1-29.

Blair, R.J.R. (1997). Moral reasoning in the child with psychopathic tendencies. *Personality and Individual Differences, 22*, 731-739.

Blair, R.J.R. (2001). Neuro-cognitive models of aggression, the Antisocial Personality Disorders and Psychopathy. *Journal of Neurology, Neurosurgery & Psychiatry, 71*, 727-731.

Blair, R.J.R. (2003a). Neurobiological basis of psychopathy. *British Journal of Psychiatry, 182*, 5-7.

Blair, R.J.R. (2003b). A neurocognitive model of the psychopathic individual. M.A. Ron & T.W. Robbins (Eds.), *Disorders of Brain and Mind 2* (pp. 400-420). Cambridge: Cambridge Press.

Blair, R.J.R. (2003c). Did Cain fail to represent the thoughts of Abel before he killed him? The relationship between theory of mind and aggression. B. Rappacholi & V. Slaughter (Eds.), *Individual differences in theory of mind: Implications for typical and atypical development. (pp. 143-170).* New York, Psychology Press.

Blair, R.J.R. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior. *Brain and Cognition, 55*(1), 198-208.

Blair, R.J.R., & Cipolotti, L. (2000). Impaired social response-reversal: A case of "acquired sociopathy". *Brain, 123*, 1122-1141.

Blair, R.J.R., Colledge, E., & Mitchell, D.G.V. (2001a). Somatic markers and response reversal: is there orbitofrontal cortex dysfunction in boys with psychopathic tendencies? *Journal of Abnormal Child Psychology, 29*(6), 499-511.

Blair, R.J.R., Colledge, E., Murray, L., & Mitchell, D.G.V. (2001b). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology, 29*(6), 491-498.

Blair, R.J.R. & Curran, H.V. (1999). Selective impairment in the recognition of anger induced by diazepam. *Psychopharmacology, 147,* 335-338.

Blair, R.J.R., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology, 34*, 192-198.

Blair, R.J.R., Mitchell, D.G.V., Leonard, A., Budhani, S., Peschardt, K.S., & Newman, C. (2004a). Passive avoidance learning in individuals with psychopathy: modulation by reward but not by punishment. *Personality and Individual Differences, 37*(6), 1179-1192.

Blair, R.J.R., Mitchell, D.G.V., Peschardt, K.S., Colledge, E., Leonard, R.A., Shine, J.H., Murray, L.K., & Perrett, D.I. (2004b). Reduced sensitivity to others' fearful expressions in psychopathic individuals. *Personality and Individual Differences, 37*(6), 1111-1122.

Blair, R.J.R., Mitchell, D.G.V., Richell, R.A., Kelly, S., Leonard, A., Newman, C., & Scott, S. (2002). Turning a deaf ear to fear: Impaired recognition of vocal affect in psychopathic individuals. *Journal of Abnormal Psychology, 111*, 682-686.

Blair, R.J.R., Morris, J.S., Frith, C.D., Perrett, D.I., & Dolan, R. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain, 122*, 883-893.

Blair, R.J.R., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., & Jones, L. (1996). Theory of Mind in the psychopath. *Journal of Forensic Psychiatry, 7,* 15-25.

Blanchard, R.J., Blanchard, D.C., & Takahashi, L.K. (1978). Pain and aggression in the cat. *Behavioral Biology, 23*(3), 291-305.

Blumer, D., & Benson, D.F. (1975). Personality changes with frontal and temporal lobe lesions. In D.F. Benson & D. Blumer (Eds.), *Psychiatric Aspects of Neurological Disease* (pp. 151-170). New York: Grune & Stratton.

Borrill, J.A., Rosen, B.K., & Summerfield, A.B. (1987). The influence of alcohol on judgment of facial expressions of emotion. *British Journal of Medical Psychology, 60*, 71-77.

Breiter, H.C., Etcoff, N.L., Whalen, P.J., Kennedy, W.A., Rauch, S.L., Buckner, R.L., Strauss, M.M., Hyman, S.E., & Rosen, B.R. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron, 17*, 875-887.

Budhani, S. (2004; in preparation). *Amygdala and Orbitofrontal Cortex Functioning in Children with Psychopathic Tendencies*. Unpublished doctoral dissertation, University College London, London.

Brothers, L., & Ring, B. (1992). Aneuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience, 4,* 107-118.

Burgess, P.W., & Shallice, T. (1993). Guessing, rule detection and frontal lobe lesions. *Cognitive Neuropsychology.*

Burgess, P.W., & Wood, R.L. (1990). Neuropsychology of behaviour disorders following brain injury. R.L. Wood (Ed.), *Neurobehavioural Sequelae of Traumatic Brain Injury*. London, Taylor & Francis: 110-133.

Burns, L.H., Everitt, B.J., & Robbins, T.W. (1999). Effects of excitotoxic lesions of the basolateral amygdala on conditional discrimination learning with primary and conditioned reinforcement. *Behavioural Brain Research, 100*, 123-133.

Butter, C.M. (1969). Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in Macaca mulatta. *Physiology & Behavior, 4*(2), 163-171.

Carmichael, S. T., & Price, J. L. (1995). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology, 363*(4), 642-664.

Chow, T.W., & Cummings, J.L. (2000). The amygdala and alzheimer's disease. In J.P. Aggleton (Ed.), *The Amygdala* (pp. 655-680). Oxford: Oxford University Press.

Clark, L., Cools, R., & Robbins, T.W. (2004). The neuropsychology of ventral prefrontal cortex: decision-making and reversal learning. *Brain and Cognition, 55*(1), 41-53.

Clarke, H.F., Dalley, J.W., Crofts, H.S., Robbins, T.W., & Roberts, A.C. (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science, 304*(5672), 878-80.

Cleckley, H. (1967). *The mask of sanity.* St Louis, MO, Mosby.

Cools, R., Clark, L., Owen, A.M., & Robbins, T. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *The Journal of Neuroscience, 22*(11), 4563-4567.

Cools, R., Clark, L., & Robbins, T.W. (2004). Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience, 24*(5), 1129-1135.

Corcoran, R., Cahill, C. & Frith, C.D. (1997). The appreciation of visual jokes in people with schizophrenia: A study of "mentalizing" ability. *Schizophrenia Research, 24,* 319-327.

Corcoran, R., Mercer, G., & Frith, C.D. (1995). Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophrenia Research, 17*(1), 5-13.

Cornell, D.G., Warren, J., Hawk, G., Stafford, E., Oram, G., & Pine, D. (1996). Psychopathy in instrumental and reactive violent offenders. *Journal of Consulting and Clinical Psychology, 64*, 783-790.

Coull, J.T., Middleton, H.C., Young, A.H., Park, S.B., McShane, R.H., Cowen, P.J., & Robbins, T.W. (1995). Clonidine and diazepam have differential effects on tests of attention and learning. *Psychopharmacology, 120*(3): 322-332.

Critchley, H.D., & Rolls, E.T. (1996). Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *Journal of Neurophysiolology, 75*(4), 1673-1686.

Damasio, A.R. (1994). *Descartes' error: Emotion, rationality and the human brain.* New York: Putnam (Grosset Books).

Damasio, A. (1998). The somatic marker hypothesis and the possible functions of the prefrontal cortex. A.C. Roberts, T.W. Robbins & L. Weiskrantz (Eds.), *The Prefrontal Cortex* (pp. 36-50). New York, NY, USA: Oxford University Press.

Damasio, A.R., Tranel, D., & Damasio, H. (1990). Individuals with sociopathic behaviour caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research, 41*, 81-94.

Damasio, A.R., Tranel, D., & Damasio, H.C. (1991). Somatic markers and the guidance of behavior: Theory and preliminary testing. In H.S. Levin, H.M. Eisenberg &

A.L. Benton (Eds.), *Frontal Lobe Function and Dysfunction* (pp. 217-229). New York: Oxford University Press.

Davis, M. (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. J.P. Aggleton (Ed.), *The Amygdala: A functional analysis*. Oxford: Oxford University Press: 289-310.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Dewey, M. (1991). Living with Asperger's syndrome. U. Frith (Ed.) *Autism and Asperger's Syndrome*. Cambridge, Cambridge University Press.

Dias, R., Robbins, T.W., & Roberts, A.C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*, 69-72.

Dolan, R.J., Fletcher, P., Morris, J., Kapur, N., Deakin, J.F., & Frith, C.D. (1996). Neural activation during covert processing of positive emotional facial expressions. *Neuroimage, 4*, 194-200.

Dolan, R.J., & Vuilleumier, P. (2003). Amygdala automaticity in emotional processing. *Annals of the NewYork Academy of Sciences, 985*, 348-355.

Drevets, W.C., Videen, T.O., Price, J.L., Preskorn, S.H., Carmichael, S.T., & Raichle, M.E. (1992). A functional anatomical study of unipolar depression. *Journal of Neuroscience, 12*(9), 3628-3641.

Duncan, J. (1998). Converging levels of analysis in the cognitive neuroscience of visual attention. *Philosophical Transactions of the Royal Society, 353*, 1307-1317.

Elliott, R. & Dolan, R.J. (1999). Differential neural responses during performance of matching and nonmatching to sample tasks at two delay intervals. *Journal of Neuroscience, 19*(12), 5066-5073.

Elliott, R., Dolan, R.J., & Frith, C.D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cerebral Cortex 10*(3): 308-17.

Elliott, R., Rees, G., & Dolan, R.J. (1999). Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia 37*(4), 403-11.

Emery, N.J., Capitanio, J.P., Mason, W.A., Machado, C.J., Mendoza, S.P., & Amaral, D.G. (2001). The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (Macaca mulatta). *Behavioral Neuroscience, 115*(3), 515-544.

Ernst, M., Bolla, K., Mouratidis, M., Contoreggi, C., Matochik, J.A., Kurian, V., Cadet, J., Kimes, A.S., & London, E.D. (2002). Decision-making in a risk-taking task: A PET study. *Neuropsychopharmacology, 26*(5), 682-691.

Eslinger, P.J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology 39*, 193-199.

Estes, W.K. & Skinner, B.F. (1941). Some quantitative properties of anxiety. *Journal of Experimental Psychology 29*, 390-400.

Everitt, B.J., Cardinal, R.N., Parkinson, J.A., & Robbins, T.W. (2003). Appetitive behavior: impact of amygdala-dependent mechanisms of emotional learning. *Annals of the New York Academy of Science, 985*, 233-250.

Fellows, L.K., & Farah, M.J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain, 126*, 1830-1837.

Fine, C., & Blair, R.J.R. (2000). Mini review: The cognitive and emotional effects of amygdala damage. *Neurocase, 6*, 435-450.

Fine, C.J. (2000). *Expectation Violations and Emotional Learning.* Unpublished doctoral dissertation, University College London, London.

Fine, C., J. Lumsden, & Blair, R.J.R. (2001). Dissociation between "theory of mind" and executive functions in a patient with early left amygdala damage. *Brain, 124*, 287-98.

Fisher, L., & Blair, R.J.R. (1998). Cognitive impairment and its relationship to psychopathic tendencies in children with emotional and behavioural difficulties. *Journal of Abnormal Child Psychology, 26*, 511-519.

Flor, H., Birbaumer, N., Hermann, C., Ziegler, S., & Patrick, C.J. (2002). Aversive Pavlovian conditioning in psychopaths: peripheral and central correlates. *Psychophysiology, 39*(4), 505-518.

Fowles, D.C. (1988). Psychophysiology and Psychopathology: A motivational approach. *Psychophysiology 25*(4), 373-391.

Frith, C.D., & Corcoran, R. (1996). Exploring "theory of mind" in people with schizophrenia. *Psychological Medicine, 26*(3): 521-30.

Fudge, J.L., Powers, J.M., Haber, S.N., & Caine, E.D. (1998). Considering the role of the amygdala in psychotic illness: a clinicopathological correlation. *Journal of Neuropsychiatry and Clinical Neurosience, 10*(4), 383-394.

Fuster, J.M. (1997). *The Prefrontal Cortex Anatomy, Physiology, and Neuropsychology of the Frontal Lobe* (3rd ed.). Philadelphia: Lippincott-Raven.

Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., & Frith, C.D. (2000). Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia 38*(1), 11-21.

Garavan, H., Pendergrass, J.C., Ross, T.J., Stein, E.A., & Risinger, R.C. (2001). Amygdala response to both positively and negatively valenced stimuli. *Neuroreport, 12*(12): 2779-83.

Gilboa, A., Shalev, A.Y., Laor, L., Lester, H., Louzoun, Y., Chisin, R., & Bonne, O. (2004). Functional connectivity of the prefrontal cortex and the amygdala in posttraumatic stress disorder. *Biological Psychiatry, 55*(3), 263-272

Grafman, J., Schwab, K., Warden, D., Pridgen, B.S., & Brown, H.R. (1996). Frontal lobe injuries, violence, and aggression: A report of the Vietnam head injury study. *Neurology, 46*, 1231-1238.

Grant, S., Contoreggi, C., & London, E.D. (2000). Drug abusers show impaired performance in a laboratory test of decision-making. *Neuropsychologia, 38*, 1180-1187.

Grattan, L.M., Bloomer, R.H., Archambault, F.X., & Eslinger, P.J. (1994). Cognitive flexibility and empathy after frontal lobe lesion. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology 7,* 251-259.

Gray, J.A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge: University of Cambridge Press.

Gregg, T.R., & Siegel, A. (2001). Brain structures and neurotransmitters regulating aggression in cats: implications for human aggression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 25*(1), 91-140.

Happé, F.G.E. (1994). An advanced test of Theory of Mind: Understanding of story characters" thoughts and feelings in able autistic, mentally handicapped, and normal children and adults. *Journal of autism and developmental disorders, 24,* 129-154.

Hare, R.D. (1984). Performance of psychopaths in cognitive tasks related to frontal lobe function. *Journal of Abnormal Psychology, 93*, 133-140.

Hare, R.D. (1991). *The Hare Psychopathy Checklist-Revised.* Toronto, Ontario: Multi-Health Systems.

Hare, R.D. (1998). Psychopathy, affect and behavior. In D.J. Cooke, A.E. Forth & R.D. Hare (Eds.), *Psychopathy: Theory, Research and Implications for Society* (pp. 81-105). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hare, R.D. (2003). *Hare Psychopathy Checklist-Revised (PCL-R).* (2nd Ed). Toronto, MHS.

Hare, R.D., & Quinn, M.J. (1971). Psychopathy and autonomic conditioning. *Journal of Abnormal Psychology, 77*, 223-235.

Harmer, C.J., Bhagwagar, Z., Cowen, P.J., & Goodwin, G.W. (2001). Acute adiminstration of citalopram in healthy volunteers facilitates recognition of happiness and fear. *Journal of Psychopharmacology, 15*, (Supplement), A16.

Harlow, J.M. (1868). Recovery from the passage of an iron bar through the head. *Publications of the Massachusetts Medical Society, 2*, 327-347.

Harris, G.T., Rice, M.E., & Cormier, C.A. (1991). Psychopathy and violent recidivism. *Law & Human Behavior, 15*, 625-637.

Hayman, L.A., Rexer, J.L., Pavol, M.A., Strite, D., & Meyers, C.A. (1998). Kluver-Bucy syndrome after bilateral selective damage of amygdala and its cortical connections. *Journal of Neuropsychiatry and Clinical Neuroscience, 10*(3), 354-358.

Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R., & Polkey, C.E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain, 126*, 1691-1712.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., & Polkey, C.E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience, 16*(3), 463-478.

Hornak, J., Rolls, E.T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal damage. *Neuropsychologia, 34*, 247-261.

Insausti, R., Amaral, D.G., & Cowan, W.M. (1987). The entorhinal cortex of the monkey: II. Cortical afferents. *Journal of Comparative Neurology, 264*(3), 356-395.

LaBar, K.S., Gatenby, J.C., Gore, J.C., LeDoux, J.E., & Phelps, E.A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron, 20*(5), 937-945.

LaBar, K.S., LeDoux, J.E., Spencer, D.D., & Phelps, E.A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *Journal of Neuroscience, 15*, 6846-6855.

LaPierre, D., Braun, C.M.J., & Hodgins, S. (1995). Ventral frontal deficits in psychopathy: Neuropsychological test findings. *Neuropsychologia, 33*, 139-151.

Levenston, G.K., Patrick, C.J., Bradley, M.M., & Lang, P.J. (2000). The psychopath as observer: Emotion and attention in picture processing. *Journal of Abnormal Psychology, 109*, 373-386.

Lykken, D.T. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology, 55*, 6-10.

Lykken, D.T. (1995). *The Antisocial Personalities*. Hillsdale, New Jersey: Erlbaum.

Kanwisher, N., Stanley, D., Harris, A. (2000). The fusiform face area is selective for faces not animals. *Neuroreport 18,* 183-187.

Kawasaki, H., Adolphs, R., Kaufman, O., Damasio, H., Damasio, A.R., Granner, M., Bakken, H., Hori, T., & Howard, M.A. (2001). Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience, 4*(1), 15-16.

Kesler/West, M.L., Andersen, A.H., Smith, C.D., Avison, M.J., Davis, C.E., Kryscio, R.J., & Blonder, L.X. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research, 11*(2), 213-226.

Kiehl, K.A., Smith, A.M., Hare, R.D., Mendrek, A., Forster, B.B., Brink, J., & Liddle, P.F. (2001). Limbic Abnormalities in Affective Processing By Criminal

Psychopaths as Revealed by Functional Magnetic Resonance Imaging. *Biological Psychiatry, 50*, 677-684.

Killcross, S., Robbins, T.W., & Everitt, B.J. (1997). Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature, 388,* 377-380.

Kringelbach, M.L., & Rolls, E.T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. Progress in Neurobiology, 72, 341-372.

Kluver, H., & Bucy, P.C. (1997). Preliminary analysis of functions of the temporal lobes in monkeys. 1939. *Journal of Neuropsychiatry and Clinical Neuroscience, 9*(4), 606-620.

Kosson, D.S., Suchy, Y., Mayer, A.R., & Libby, J. (2002). Facial affect recognition in criminal psychopaths. *Emotion, 2*(4), 398-411.

Lang, P.J., & Greenwald, M.K. (1988). *The international affective picture system standardization procedure and initial group results for affective judgements: Technical reports 1A & 1B*. Gainesville, Center for Research in Psychophysiology, University of Florida.

LeDoux, J. (1998). *The Emotional Brain*. New York, Weidenfeld & Nicolson.

LeDoux, J. E., Cicchetti, P., Xagoraris, A., & Romanski, L.M. (1990). The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. *Journal of Neuroscience, 10*(4), 1062-1069.

Liberzon, I., Phan, K.L., Decker, L.R., & Taylor, S.F. (2003). Extended amygdala and emotional salience: a PET activation study of positive and negative affect. *Neuropsychopharmacology 28*(4), 726-733.

Lorenz, A. R. and J. P. Newman (2002). Deficient response modulation and emotion processing in low-anxious Caucasian psychopathic offenders: results from a lexical decision task. *Emotion, 2*(2), 91-104.

Malkova, L., Gaffan, D., & Murray, E.A. (1997a). Excitotoxic lesions of the amygdala fail to produce impairment in visual learning of auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. *The Journal of Neuroscience, 17*(15), 6011-6020.

Malkova, L., Mishkin, M., Suomi, S., & Bachelvalier, J. (1997b). Socioemotional behavior in adult rhesus monkeys after early versus late lesions of the medial temporal lobe. *Annals of the New York Academy of Sciences, 807*, 538-540.

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain, 125*, 624-639.

Mavaddat, N., Kirkpatrick, P.J., Rogers, R.D., & Sahakian, B.J. (2000). Deficits in decision-making in patients with aneurysms of the anterior communicating artery. *Brain, 123*, 2109-2117.

McKenna, P., & Warrington, E.K. (1980). Testing for nominal dysphasia. *Journal of Neurology,Neurosurgery,and Psychiatry, 43*, 781-788.

Mitchell, D.G.V., Colledge, E., Leonard, A., & Blair, R.J.R. (2002). Risky decisions and response-reversal: Is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? *Neuropsychologia, 40*, 2013-2022.

Morris, J.S., Frith, C.D., Perrett, D.I., Rowland, D., Young, A.W., Calder, A.J., & Dolan, R.J. (1996). A differential response in the human amygdala to fearful and happy facial expressions. *Nature, 383*, 812-815.

Morris, J.S., Ohman, A., & Dolan, R.J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Science, 96,* 1680-1685.

Morris, J.S., Scott, S.K., & Dolan, R.J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia, 37*(10), 1155-1163.

Nelson, H.E., & Willison, J. (1991). *The National Adult Reading Test* (2nd ed.). Windsor (UK): NFER-Nelson.

Newman, J.P. (1998). Psychopathic behaviour: An information processing perspective. In D.J. Cooke, A.E. Forth & R.D. Hare (Eds.), *Psychopathy: Theory, Research and Implications for Society* (pp. 81-105). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Newman, J.P., & Kosson, D.S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology, 95*, 252-256.

Newman, J.P., Patterson, C.M., & Kosson, D.S. (1987). Response perseveration in psychopaths. *Journal of Abnormal Psychology, 96*, 145-148.

Newman, J.P., & Schmitt, W.A. (1998). Passive avoidance in psychopathic offenders: a replication and extension. *Journal of Abnormal Psychology, 107*, 527-532.

Newman, J.P., Schmitt, W.A., & Voss, W.D. (1997). The impact of motivationally neutral cues on psychopathic individuals: Assessing the generality of the response modulation hypothesis. *Journal of Abnormal Psychology 106*, 563-575.

Newman, J.P., Wallace, J.F., Schmitt, W.A., & Arnett, P.A. (1997). Behavioral inhibition system functioning in anxious, impulsive and psychopathic individuals. *Personality and Individual Differences, 23*(4): 583-592.

Newman, J.P., Widom, C.S., & Nathan, S. (1985). Passive avoidance in syndromes of disinhibition: Psychopathy and extraversion. *Journal of Personality and Social Psychology, 48*, 1316-1327.

Nutt, D.J., & Malizia, A.L. (2004). Structural and functional brain changes in posttraumatic stress disorder. *Journal of Clinical Psychiatry, 65*(suppl 1), 11-17.

O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R.J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of Neuroscience, 23*(21), 7931-7939.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., & Andrews, C. (2001a). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*(1), 95-102.

O'Doherty, J., Rolls, E.T., Francis, S., Bowtell, R., & McGlone, F. (2001b). Representation of pleasant and aversive taste in the human brain. *Journal of Neurophysiology, 85*(3), 1315-21.

O'Doherty, J., Rolls, E.T., Francis, S., Bowtell, R., McGlone, F., Kobal, G., Renner, B., & Ahne, G. (2000). Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport, 11*(4), 893-7.

O'Doherty, J.P., Deichmann, R., Critchley, H.D., & Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron, 33*(5), 815-26.

Ogloff, J.R.P., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences and the Law, 8*, 181-190.

Ongur, D., & Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex, 10*(3), 206-219.

Panksepp, J. (1971). Aggression elicited by electrical stimulation of the hypothalamus in albino rats. *Physiology and Behavior, 6*(4), 321-329.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Pastor, M.C., Molto, J., Vila, J., & Lang, P.J. (2003). Startle reflex modulation, affective ratings and autonomic reactivity in incarcerated Spanish psychopaths. *Psychophysiology, 40*, 934-938.

Patrick, C.J. (1994). Emotion and psychopathy: Startling new insights. *Psychophysiology, 31*, 319-330.

Patrick, C.J., Bradley, M.M., & Lang, P.J. (1993). Emotion in the criminal psychopath: Startle reflex modulation. *Journal of Abnormal Psychology, 102*, 82-92.

Peschardt, K. S., Leonard, R., Richell, R.A., & Blair, R.J.R. (2004). Differential Stimulus-Reward and Stimulus-Punishment Learning in Individuals with Psychopathy. Manuscript submitted for publication-a.

Peschardt, K. S., Richell, R.A., & Blair, R.J.R. (2004). Primed up for positive, but not negative, words: Affective priming in individuals with psychopathy. Manuscript submitted for publication-b.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L.G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Science, 99*(17), 11458-63.

Petrides, M. (1982). Motor conditional associative-learning after selective prefrontal lesions in the monkey. *Behavioural Brain Research, 5*(4), 407-413.

Petrides, M. (1985a). Deficits in non-spatial conditional associative learning after periarcuate lesions in the monkey. *Behavioural Brain Research 16*(2-3), 95-101.

Petrides, M. (1985b). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 23*(5), 601-614.

Petrides, M. (1990). Nonspatial conditional learning impaired in patients with unilateral frontal but not unilateral temporal lobe excisions. *Neuropsychologia 28*(2), 137-149.

Petrides, M. (1997). Visuo-motor conditional associative learning after frontal and temporal lesions in the human brain. *Neuropsychologia 35*(7), 989-97.

Petrides, M., & Pandya, D.N. (2001). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience, 16*(2), 291-310.

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., Williams, S.C.R., Bullmore, E.T., Brammer, M., & Gray, J.A. (1998).

Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society of London B, 265*(1408), 1809-1817.

Phillips, M.L., Young, A.W., Senior, C., Brammer, M., Andrews, C., Calder, A.J., Bullmore, E.T., Perrett, D.I., Rowland, D., Williams, S.C.R., Gray, J.A., & David, A.S. (1997). A specified neural substrate for perceiving facial expressions of disgust. *Nature, 389*, 495-498.

Prather, M.D., Lavenex, P., Mauldin-Jourdain, M.L., Mason, W.A., Capitanio, J.P., Mendoza, S.P., & Amaral, D.G. (2001). Increased social fear and decreased fear of objects in monkeys with neonatal amygdala lesions. *Neuroscience, 106*(4), 653-658.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences, 1*(4): 515-526.

Price, J.L. (2003). Comparative aspects of amygdala connectivity. *Annals of the New York Academy of Science, 985*, 50-58.

Raine, A., Buchsbaum, M.S., & LaCasse, L. (1997). Brain abnormalities in murderers indicated by positron emission tomography. *Biological Psychiatry, 42*, 495-508.

Raine, A., Buchsbaum, M.S., Stanley, J., Lottenberg, S., Abel, L., & Stoddard, J. (1994). Selective reductions in prefrontal glucose metabolism in murderers. *Biological Psychiatry, 15*, 365-373.

Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry, 57*, 119-127.

Raine, A., Meloy, J. R., Birhle, S., Stoddard, J., LaCasse, L., & Buchsbaum, M.S. (1998). Reduced prefrontal and increased subcortical brain functioning assessed using positron emission tomography in predatory and affective murderers. *Behaviour Science and Law, 16*, 319-332.

Ramamurthi, B. (1988). Stereotactic operation in behaviour disorders. Amygdalotomy and hypothalamotomy. *Acta Neurochirgica, 44*, 152-157.

Rapcsak, S.Z., Galper, S.R., Comer, J.F., Reminger, S.L., Nielsen, L., Kaszniak, A.W., Verfaelli, M., Laguna, J.F., Labiner, D.M., & Cohen, R.A. (2000). Fear recognition deficits after focal brain damage: A cautionary note. *Neurology, 54*, 575-581.

Raven, J. (1965; 1994). Colored Progressive Matrices Sets A, Ab, B. Oxford: Oxford Psychologists Press Ltd.

Rice, M.E., Harris, G.T., & Cormier, C.A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior, 16*, 399-412.

Richell, R.A., Mitchell, D.G.V., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R.J. (2003). Theory of mind and psychopathy: can psychopathic individuals read the "language of the eyes"? *Neuropsychologia 41*(5), 523-6.

Rilling, J. K., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., & Cohen, J.D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage 22*(4), 1694-1703.

Rogers, R.D., Andrews, T.C., Grasby, P.M., Brooks, D.J., & Robbins, T.W. (2000). Contrasting cortical and subcortical activations produced by attentional-set

shifting and reversal learning in humans. *Journal of Cognitive Neuroscience, 12*(1), 142-62.

Rogers, R.D., Everitt, B.J., Baldacchino, A., Blackshaw, A.J., Swainson, R., Wynne, K., Baker, N.B., Hunter, J., Carthy, T., Booker, E., London, M., Deakin, J.F.W., Sahakian, B.J., & Robbins, T.W. (1999a). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology, 20*(4), 322-339.

Rogers, R.D., Owen, A.M., Middleton, H.C., Williams, E.J., Pickard, J.D., Sahakian, B.J., & Robbins, T.W. (1999b). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience, 19*(20), 9029-9038.

Rolls, E.T. (1996). The orbitofrontal cortex. *Philosophical Transcripts of the Royal Society, 351*, 1433-1443.

Rolls, E.T. (1999). *The Brain and Emotion.* Oxford: Oxford University Press.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10*, 284-294.

Rolls, E.T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition, 55*(1), 11-29.

Rolls, E.T., & Baylis, L.L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *Journal of Neuroscience, 14*(9), 5437-5452.

Rolls, E.T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry, 57*, 1518-1524.

Rolls, E.T., Yaxley, S., & Sienkiewicz, Z.J. (1990). Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *Journal of Neurophysiology, 64*(4), 1055-1066.

Romanski, L.M., & Goldman-Rakic, P.S. (2002). An auditory domain in primate prefrontal cortex. *Nature Neuroscience, 5*(1), 15-16.

Sabbagh, M.A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: implications for autism. *Brain and Cognition 55*(1), 209-219.

Salgado, J.V., Vidal, M., Oberling, P., Graeff, F.G., Danion, J.M., & Sandner, G. (2000). Associative learning and latent in hibition in a conditioned suppression paradigm in humans. *Behavioural Brain Research 117*, 53-60.

Schindler, C.W., Thorndike, E.B., Ma, J.D., & Goldberg, S.R. (1999). Conditioned suppression with cocaine as the unconditioned stimulus. *Pharmacology Biochemistry and Behavior 65*(1), 83-89.

Schmitt, W.A., Brinkley, C.A., & Newman, J.P. (1999). Testing Damasio's somatic marker hypothesis with psychopathic individuals: Risk takers or risk averse? *Journal of Abnormal Psychology, 108*, 538-543.

Scott, S.K., Young, A.W., Calder, A.J., Hellawell, D.H., Aggleton, J.P., & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature, 385*, 254-257.

Shallice, T., & Evans, M.E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex, 14*, 294-303.

Sprengelmeyer, R., Rausch, M., Eysel, U.T., & Przuntek, H. (1998). Neural structures associated with the recognition of facial basic emotions. *Proceedings of the Royal Society of London, 265*, 1927-1931.

Sprengelmeyer, R., Young, A.W., Schroeder, U., Grossenbacher, P.G., Federlein, J., Buttner, T. (1999). Knowing no fear. *Proceedings Royal Society of London B, 266*(1437), 2451-2456.

Siebert, M., Markowitsch, H.J., & Bartel, P. (2003). Amygdala, affect and cognition: evidence from 10 patients with Urbach-Wiethe disease. *Brain, 126,* 2627-2637.

Simpson, J.R., D. Ongur, Akbudak, E., Conturo, T.E., Ollinger, J.M., Snyder, A.Z., Gusnard, D.A., & Raichle, M.E. (2000). The emotional modulation of cognitive processing: An fMRI study. *Journal of Cognitive Neuroscience, 12*(Supplement 2), 157-170.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174-215.

Stone, V. E., Baron-Cohen, S., & Knight, R.T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience 10*(5), 640-56.

Stuss, D., T., C. Gow, A., & Hetherington, C.R. (1992). "No longer Gage": Frontal lobe dysfunction and emotional changes. *Journal of Consulting and Clinical Psychology 60*: 349-359.

Swainson, R., Rogers, R.D., Sahakian, B.J., Summers, B.A., Polkey, C.E., & Robbins, T.W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia, 38*(5), 596-612

Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers, Inc.

Thorpe, S.J., Rolls, E.T., & Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Experimental Brain Research, 49*, 93-115.

Thut, G., Schultz, W., Roelcke, U., Nienhusmeier, M., Missimer, J., Maguire, R.P., & Leenders, K.L. (1997). Activation of the human brain by monetary reward. *Neuroreport, 8*(5), 1225-1228.

Tiihonen, J., Hodgins, S., Vaurio, O., Laakso, M., Repo, E., Soininen, H., Aronen, H.J., Nieminen, P., & Savolainen, L. (2000). Amygdaloid volume loss in psychopathy. *Society for Neuroscience Abstracts*, 2017.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature, 398*, 704-708.

van Elst, L.T., Woermann, F.G., Lemieux, L., Thompson, P.J., & Trimble, M.R. (2000). Affective aggression in patients with temporal lobe epilepsy: a quantitative MRI study of the amygdala. *Brain, 123,* 234-243.

van Honk, J., Hermans, E.J., Putman, P., Montagne, B., & Schutter, D.J.L.G. (1995). Defective somatic markers in sub-clinical psychopathy. *Neuroreport, 13*(8), 1027.

Veit, R., Flor, H., Erb, M., Hermann, C., Lotze, M., Grodd, W., & Birbaumer, N. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neuroscience Letters, 328*(3), 233-236.

Verona, E., Patrick, C.J., Curtin, J.J., Bradley, M.M., & Lang, P.J. (2004). Psychopathy and physiological response to emotionally evocative sounds. *Journal of Abnormal Psychology, 113*(1), 99-108.

Volkow, N.D., & Tancredi, L. (1987). Neural substrates of violent behaviour. A preliminary study with positron emission tomography. *British Journal of Psychiatry, 151*, 668-673.

Volkow, N.D., Tancredi, L.R., Grant, C., Gillespie, H., Valentine, A., Mullani, N., & Wang, G.J. (1995). Brain glucose metabolism in violent psychiatric patients: a preliminary study. *Psychiatry Research, 61*(4), 243-253.

Vuilleumier, P.J.L., Armony, J.L., Driver, J., & Dolan, R.J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron, 30*(3), 829-41.

Wallis, J.D., Dias, R., Robbins, T.W., & Roberts, A.C. (2001). Dissociable contributions of the orbitofrontal and lateral prefrontal cortex of the marmoset to performance on a detour reaching task. *European Journal of Neuroscience, 13,* 1797-1808.

Warrington, E.K. (1984). *Recognition memory test.* Windsor (UK), NFER-Nelson.

Warrington, E.K., McKenna, P., & Orpwood, L. (1998). Single word comprehension: a concrete and abstract word synonym test. *Neuropsychological Rehabilitation, 8,* 143-154.

Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Wright, C.I., & Rauch, S.L. (2001). A functional MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion, 1*(1), 70-84.

Williamson, S., Hare, R.D., & Wong, S. (1987). Violence: Criminal psychopaths and their victims. *Canadian Journal of Behavioral Science, 19*, 454-462.

Williamson, S., Harpur, T.J., & Hare, R.D. (1991). Abnormal processing of affective words by psychopaths. *Psychophysiology, 28,* 260-273.

Wong, S. (1988). Is Hare's Psychopathy Checklist reliable without the interview? *Psychological Reports, 62,* 931-934.

Yoneoka, Y., Takeda, N., Inoue, A., Ibuchi, Y., Kumagai, T., Sugai, T., Takeda, K.I., & Ueda, K. (2004). Human Kluver-Bucy syndrome following acute subdural haematoma. *Acta Neurochurgica, 146*, 1267-1270.

Zalla, T., E. Koechlin, Pietrini, P., Basso, G., Aquino, P., Sirigu, A., & Grafman, J. (2000). Differential amygdala responses to winning and losing: a functional magnetic resonance imaging study in humans. *European Journal of Neuroscience, 12*(5), 1764-1770.